NASA Contractor Report 188271

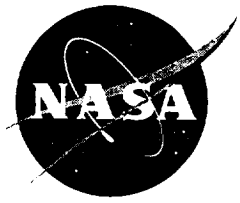
National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program - 1993

Volume 2

William A. Hyman, Editor
Texas A&M University
College Station, Texas

Stanley H. Goldstein, Editor
University Programs Office
Lyndon B. Johnson Space Center
Houston, Texas

Grant NGT 44-001-800



National Aeronautics and Space Administration
Lyndon B. Johnson Space Center

1993

Preface

The 1993 Johnson Space Center (JSC) National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program was conducted by Texas A&M University and JSC. The program at JSC, as well as the programs as other NASA centers, was funded by the Office of University Affairs, NASA Headquarters, Washington, D.C. The objectives of the program, which began nationally in 1964 and at JSC in 1965, are

1. To further the professional knowledge of qualified engineering and science faculty members

2. To stimulate an exchange of ideas between participants and NASA

3. To enrich and refresh the research and teaching activities of participants' institutions

4. To contribute to the research objectives of the NASA centers

Each faculty fellow spent at least 10 weeks at JSC engaged in a research project in collaboration with a NASA/JSC colleague. This document is a compilation of the final reports on the research projects done by the faculty fellows during the summer of 1993. Volume 1 contains reports 1 through 18 and Volume 2 contains reports 19 through 35.

Contents

Contents
(continued)

Volume 2

Contents
(concluded)

N94- 25368

DESIGN AND TESTING OF A UNIQUE RANDOMIZED GRAVITY, CONTINUOUS FLOW BIOREACTOR

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Carroll B. Lassiter, Ph.D.
Academic Rank:	Associate Professor
University and Department	University of Houston-Clear Lake Division of Biology Houston, Texas 77058
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Biomedical Laboratories
JSC Colleague:	Duane L. Pierson, Ph.D.
Date Submitted:	October 6, 1993
Contract Number:	NGT-44-001-800

(VOLUME 2)

ABSTRACT

A rotating, null gravity simulator, or Couette bioreactor has been successfully used for the culture of mammalian cells in a simulated microgravity environment. Two limited studies using Lipomyces starkeyi and Streptomyces clavuligerus were also conducted under conditions of simulated weightlessness. Although theses studies with microorganisms showed promising preliminary results, oxygen limitations presented significant limitations in studying the biochemical and cultural characteristics of these cell types. Microbial cell systems such as bacteria and yeast promise significant potential as investigative models to study the effects of micro gravity on membrane transport, as well as substrate induction of inactive enzyme systems. Additionally, the smaller size of the microorganisms should further reduce the gravity induced oscillatory particle motion and thereby improve the microgravity simulation on earth. This project focuses on the unique conceptual design, and subsequent development of a rotating bioreactor that is compatible with the culture and investigation of microgravity effects on microbial systems.The new reactor design will allow testing of highly aerobic cell types under simulated microgravity conditions. The described reactor affords a mechanism for investigating the long term effects of reduced gravity on cellular respiration, membrane transfer, ion exchange and substrate conversions. It offers the capability of dynamically altering nutrients, oxygenation, pH, carbon dioxide and substrate concentration without disturbing the microgravity simulation, or Couette flow, of the reactor. All progeny of the original cell inoculum may be acclimated to the simulated microgravity in the absence of a substrate or nutrient. The reactor has the promise of allowing scientists to probe the long term effects of weightlessness on cell interactions in plants, bacteria, yeast and fungi. The reactor is designed to have a flow field growth chamber with uniform shear stress, yet transfer high concentrations of oxygen into the culture medium. The system described allows for continuous, on line sampling for production of product without disturbing fluid and particle dynamics in the reaction chamber. It provides for the introduction of substrate, or control substances after cell adaptation to simulated microgravity has been accomplished. The reactor system provides for the non disruptive, continuous flow replacement of nutrient and removal of product. On line monitoring and control of growth conditions such as pH and nutrient status are provided. A rotating distribution valve allows for adjustment and alterations of all parameters without cessation of growth chamber rotation, thereby preserving the simulated microgravity conditions over longer periods of time.

INTRODUCTION

Clinostats have been developed and used to study higher plant growth in a simulated micro gravity for a number of years (1). Likewise, a null gravity simulator, or couette biorector has been developed and used for culture of mammalian cells in a simulated microgravity environment. Although a great deal of attention is being given to the culture of mammalian cells in a simulated micro environment,little has been done to study gravitational effect on the biochemical and cultural characteristics of microorganisms such as bacteria and fungi. Limited studies growing microorganisms in simulated weightlessness showed promising preliminary results, however oxygen limitations present problems in studying the biochemical and cultural characteristics of these microorganisms.

Cellular systems found in microbes such as bacteria, fungi and yeasts offer significant potential as investigative models for studying the effects of microgravity on membrane transport and substrate induction of inactive metabolic pathways. Additionally, because of extensive chromosomal mapping in some organisms, one can learn a great deal about mediator effects on the molecular biology of the genome.

In can be hypothesized that the smaller size of the microorganisms will reduce the gravity induced oscillatory particle motion in a couette flow bioreactor (1). It has been shown that as cell aggregates increase in size, the oscillation of the particle deviates from the ideal in a direct correlation. This suggests that the smaller size of microorganisms, and their lack of aggregation in a couette flow bioreactor would produce less gravity induced oscillation than found in mammalian cell aggregates being cultured.

The focus of this project was to develop a conceptual design and model of a rotating bioreactor that will support ideal growth conditions for microbial cell systems growing in a simulated microgravity environment. The system designed was conceived to meet, as a minimum, the following experimental provisions:

1.) Couette flow field growth chamber should present uniform shear stress.
2.) Oxygen must pass into the culture medium with a high transfer rate.
3.) System should allow for continuous, on line sampling of cell products without disturbance of fluid and particle dynamics in the cell growth chamber.
4.) The bioreactor should allow for the introduction of substrate, or control mediators, after the cell population has developed in simulated a microgravity environment.
5.) It should provide non disruptive, continuous flow replacement of nutrient.

6.) It should provide for on line adjustment of the pH of the culture media in order to maintain ideal growth conditions for a given system.

DESCRIPTION OF BIOREACTOR DESIGN

The overall flow schematic of the operating bioreactor, as conceived by the author, is shown in Figure 1.. The growth chamber of the bioreactor will be coupled to a standard bench top fermenter such as those presently being used for a wide variety of cell culture. The membrane diffusion chambers are coupled, in parallel to a roller pump, installed in the suction line, leading from the fermenter. This arrangement allows for the standard adjustment of growth conditions in the fermenter, without disturbing the cells in the growth chamber. It should be pointed out that the side diffusion chambers have been depicted much deeper, in relation to the growth chamber, than they really are. Figure 2. is an accurate representation of the size relationships between the three chambers. The thin diffusion chamber design will allow for a short residence time of media in the diffusion chamber. This will assure that the concentration gradients of nutrients, oxygen, and products will be maximized across the semipermeable membranes of the reactor. It is proposed that the bench top fermenter will be 4-6 liters in size, thereby assuring essentially infinite gradients, when interfaced with the significantly smaller bioreactor.

A detailed view of the bioreactor assembly is presented in Figure 3.. It should be pointed out that although dimensions are not presented, the reactor is designed around a 47 mm, off the shelf, general purpose, hydrophilic membrane such as those produced by a number of companies. The membranes are supported by a standard, stainless steel membrane support that is also off the shelf. The chamber provides for easy replacement of membranes and cleaning, since it is sealed by standard o-rings.

Since the bioreactor will be turning on it's central axis, it is necessary to provide for the distribution of fluid from the standard fermenter, while rotation occurs. This distribution is provided for by the rotating fluid distribution valve shown in Figure 4.. Although the valve depicted in this drawing provides for in and out distribution of a single liquid, it could easily be adapted for distribution of multiple fluids by increasing the number of grooves, or by stacking more than one valve assembly.

For compactness and simplicity, the distribution valve has been integrated into the reaction vessel compression and clamping assembly as shown in Figure 5.. This assembly provides uniform pressure on the o-ring seals in the fluid distribution valve by adjusting the compression of the tension spring. Additionally, it provides uniform sealing pressure on

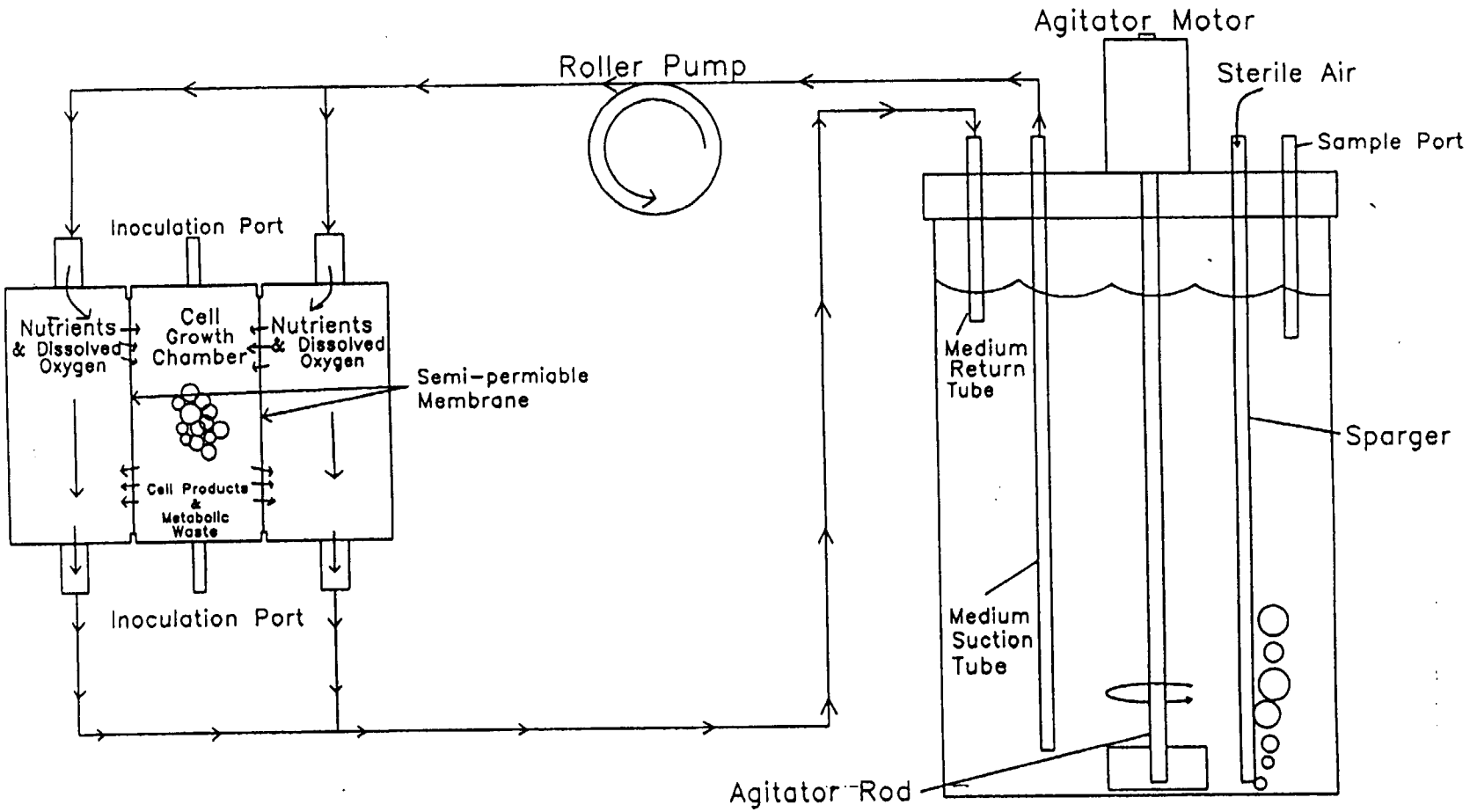


Figure 1. - General operational schematic

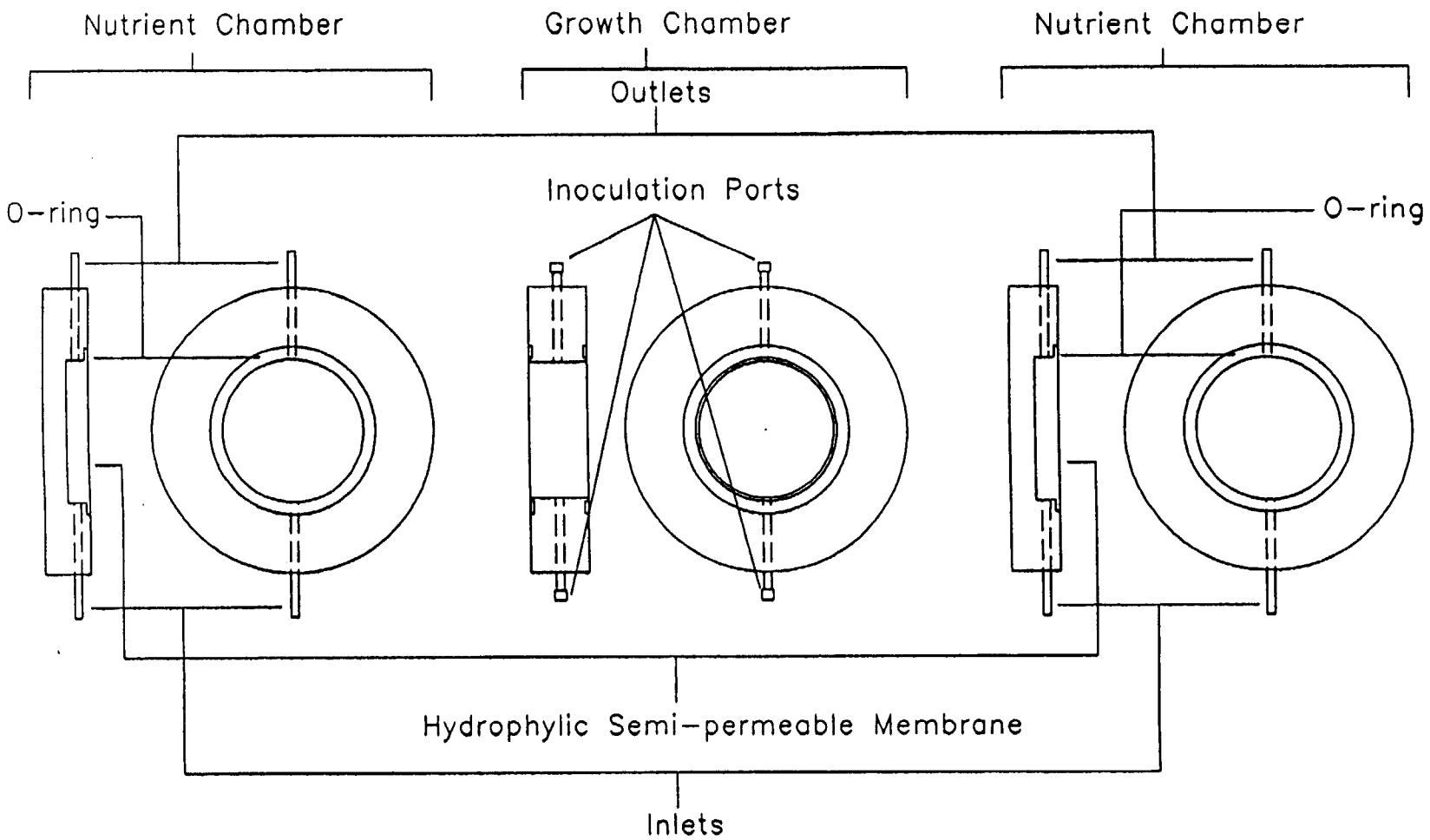


Figure 2. - Diffusion and growth chambers

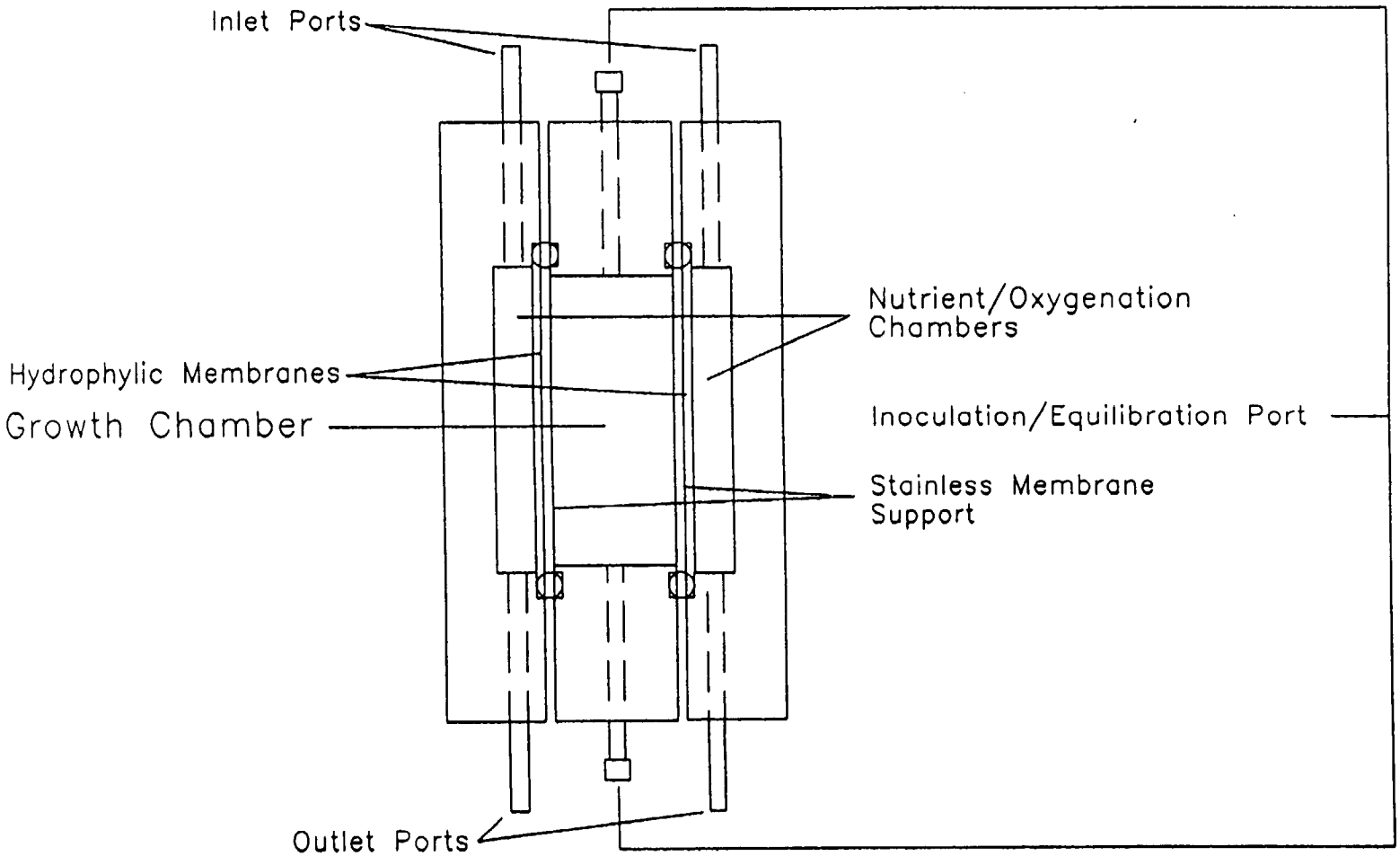


Figure 3. - Bioreactor assembly

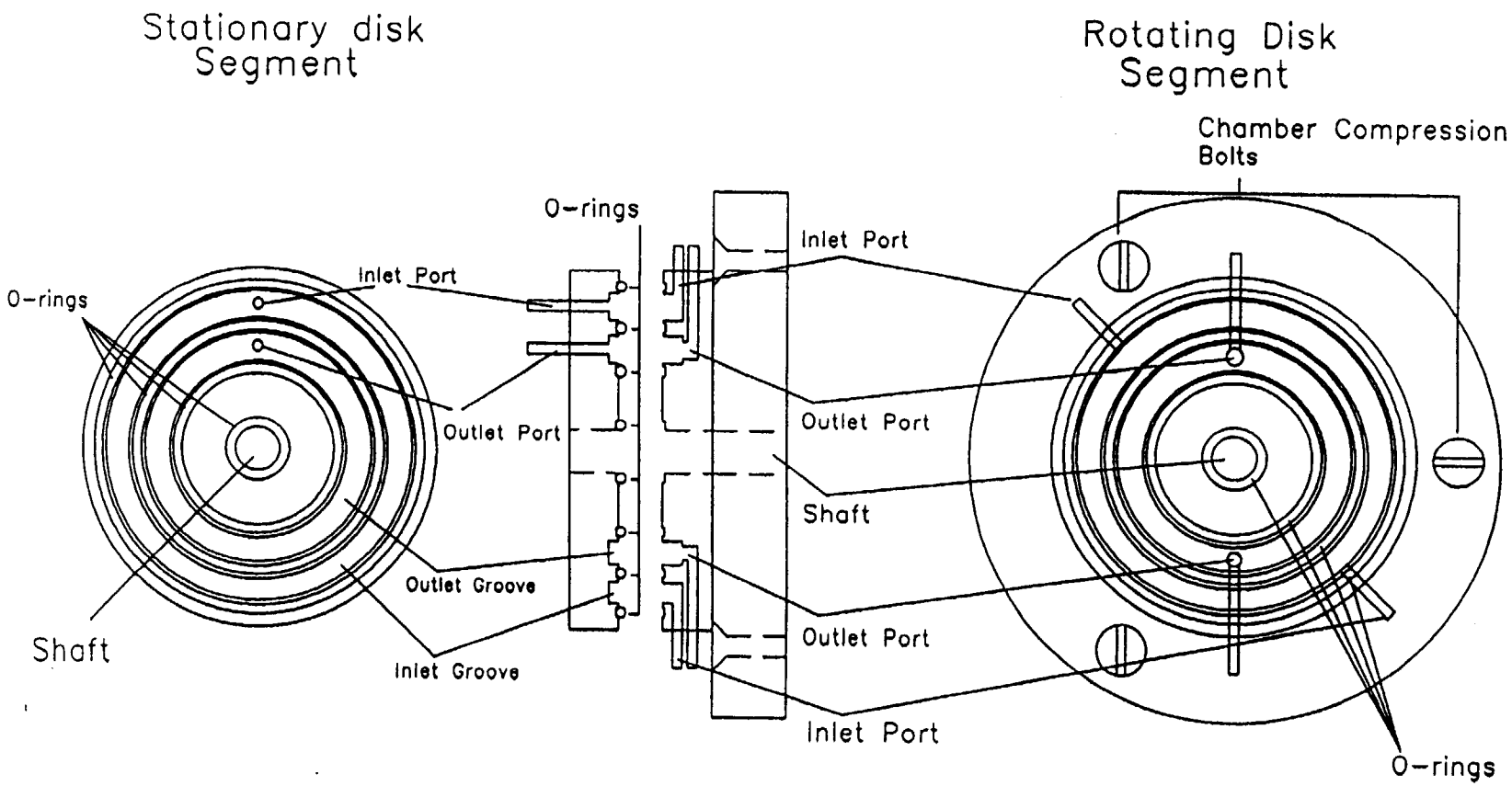


Figure 4. - Rotating fluid distribution valve

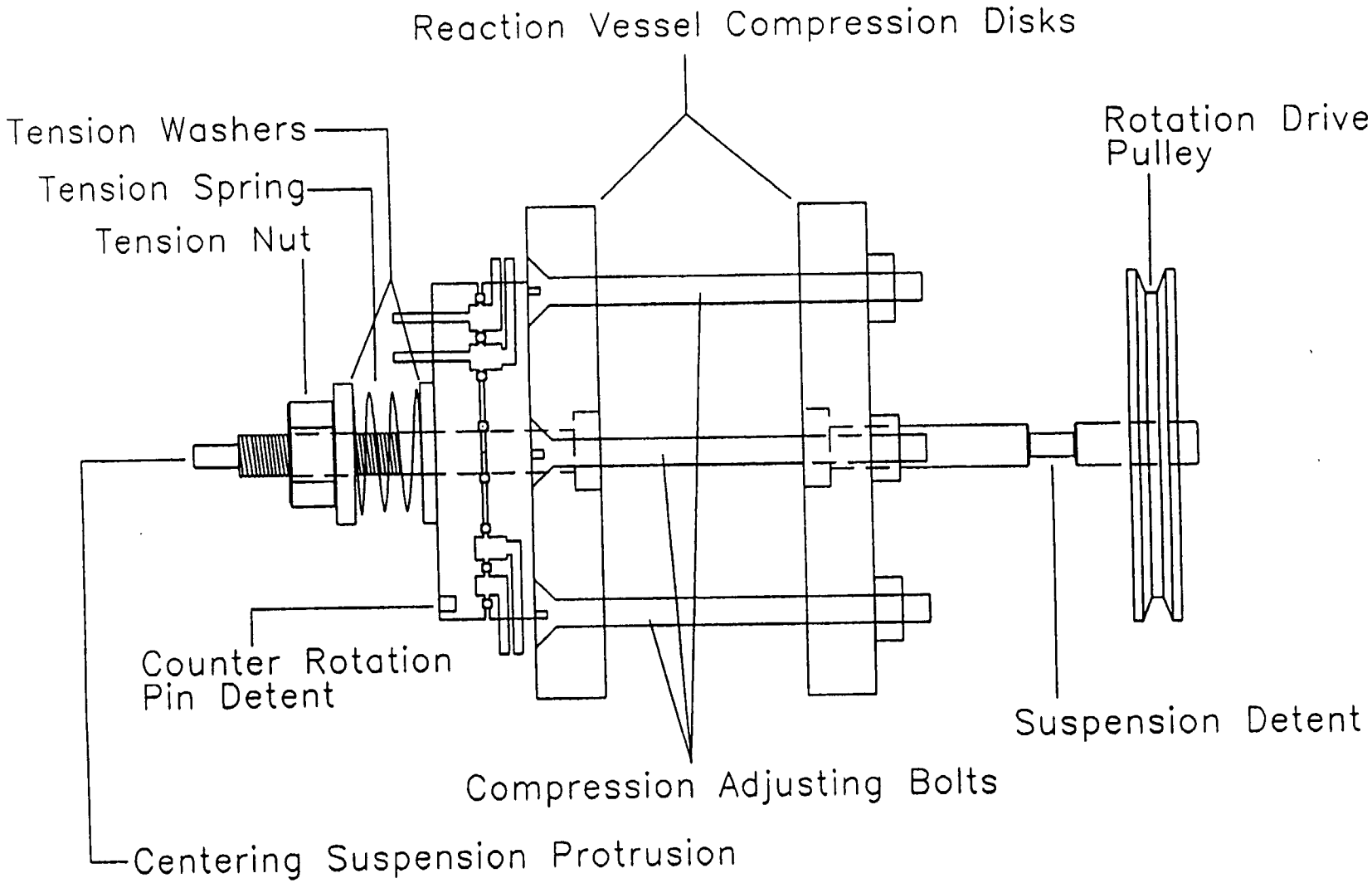


Figure 5. - Reaction vessel compression and clamping assembly

the o-ring seal in the reaction chamber by adjusting the compression of the reaction vessel compression disks. It is envisioned that the compression bolts will be tightened with a standard torque wrench, measuring inch-ounces of torque.

A schematic of the assembled, operational bioreactor is shown in Figure 6.. The integrated spacial relationships of the growth chamber, fluid distribution valve,drive pulley and clamping assembly are easily visible. For clarity the flows, to and from, the standardized bench-top fermenter are not shown, therefor it should be pointed out that the two lines emerging in the lower left corner are inlet and outlet lines coming from the fermenter.

The support cradle and drive motor for the complete chamber are shown in Figure 7.. For purposes of clarity the large rubber o-ring that is utilized as a drive belt is not shown. The drive belt itself offers sufficient tension to keep the bioreactor in the bearings of the cradle. It should be pointed out that simplicity of design has been paramount in all concepts of the proposed bioreactor. In review of the proposed design, one should find that:

1.) Critical items such as membranes, membrane supports, and o-rings would be off the shelf items.
2.) The bioreactor can be easily removed from the support cradle for manipulation, when required.
3.) The growth chamber and diffusion chambers are easily taken apart and reassembled during cleaning and sterilization.
4.) Although the conceptual model described is designed around 47 mm semipermeable membranes, the size could easily be enlarged or decreased, as required for a specific application.

SUMMARY

The proposed reactor, coupled with a bench top fermenter via a standard roller pump should allow for testing of highly aerobic cell types under simulated microgravity conditions. Since the oxygen is dissolved in the culture medium at the fermenter, the bioreactor system will require a single hydrophilic membrane type. The described reactor could afford a mechanism of investigating the long term effects of reduced gravity on cellular respiration, membrane transfer characteristic, and genetic induction by molecular mediators. It offers the capability of dynamically altering nutrients, oxygenation, pH, carbon dioxide and substrate concentrations without disturbing the couette flow of the reactor. The diffusion pressure and flows into the reaction chamber are essentially offsetting. The diffusion chambers are equal in all instances, on both sides of the growth chamber. All progeny of the original cell inoculum may be acclimated to the simulated microgravity in the absence of a substrate, nutrient

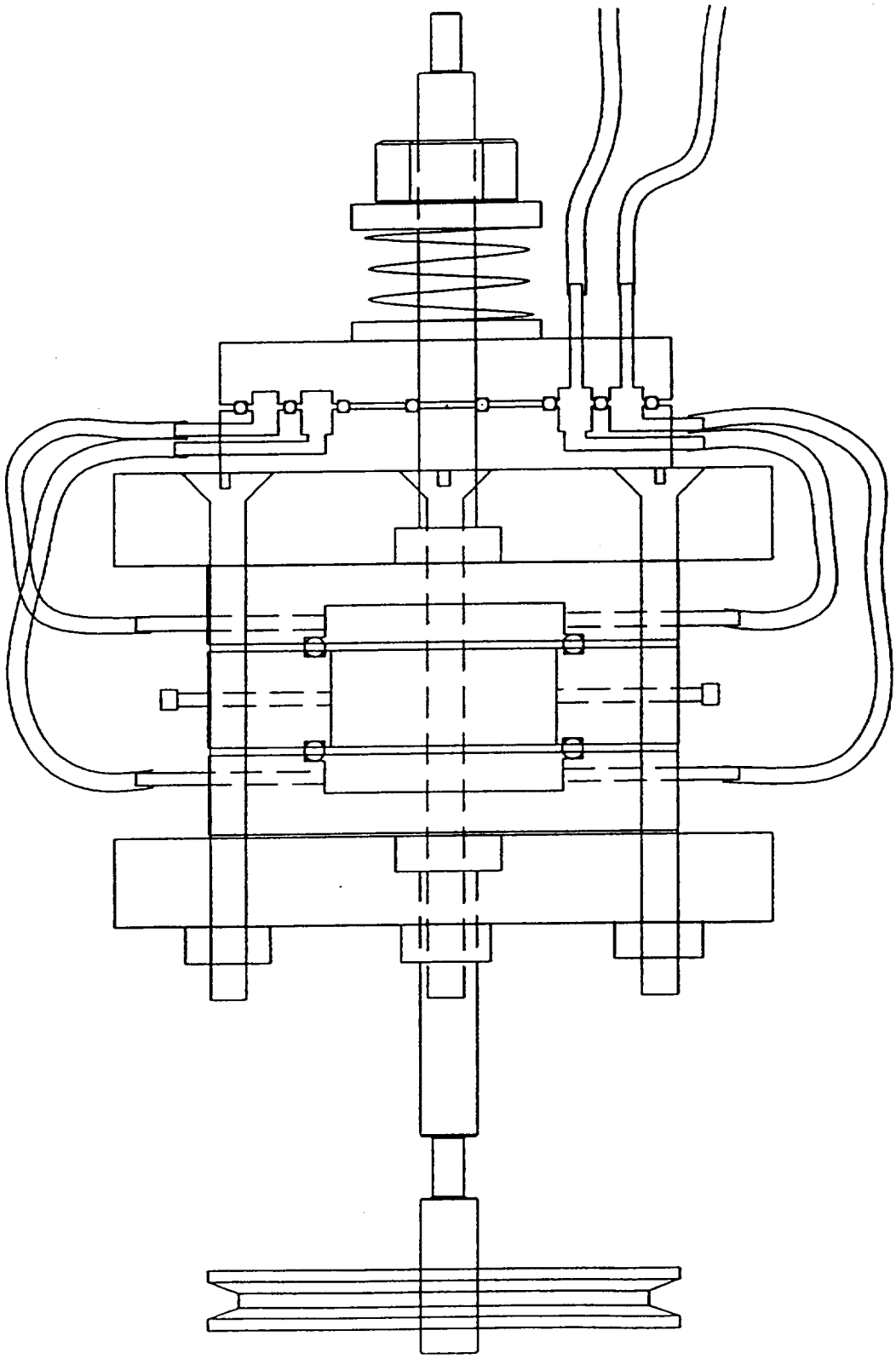


Figure 6. - Operational bioreactor assembly

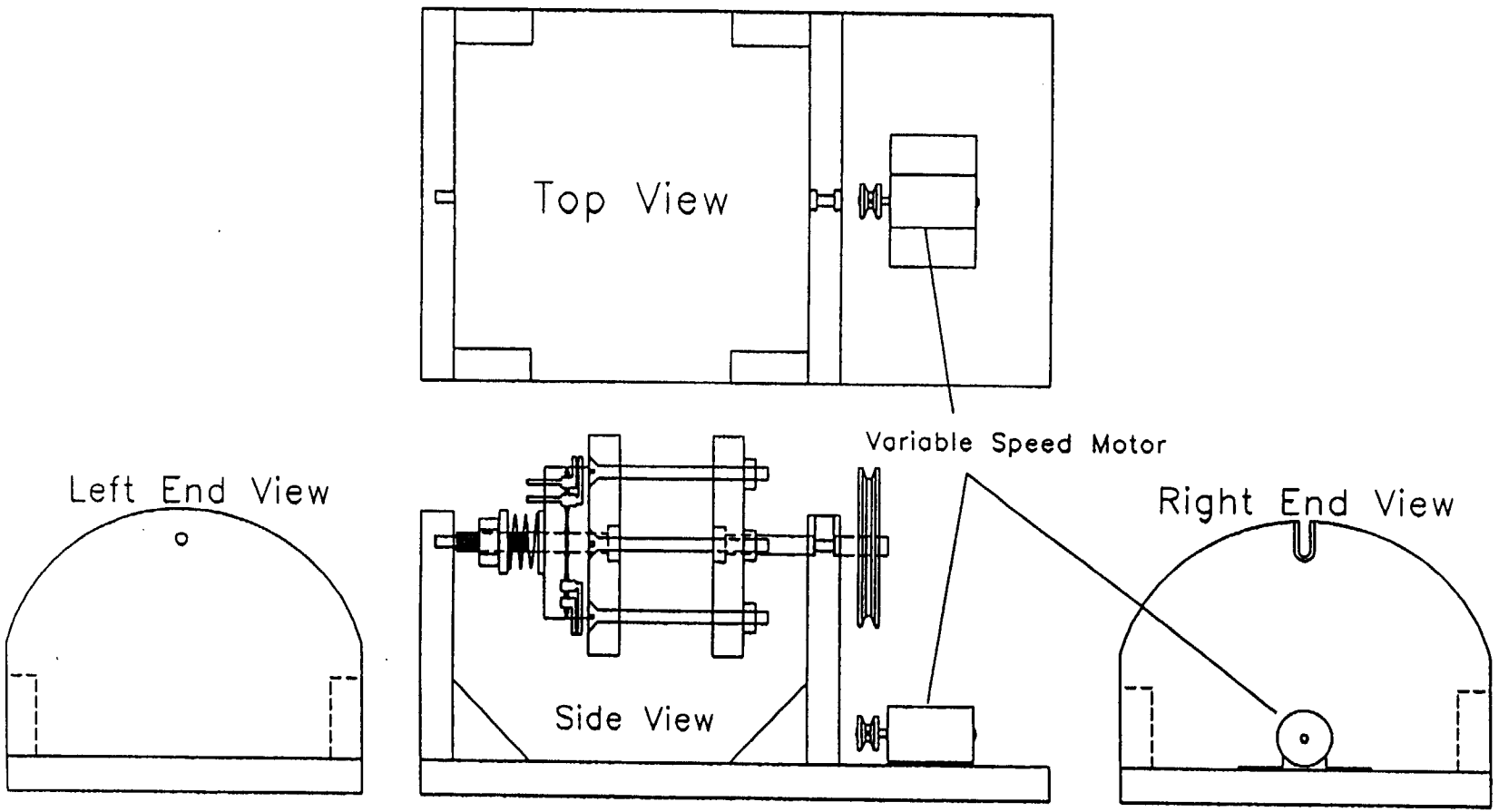


Figure 7. - Support cradle and drive motor

or control substance, whereby "batch" cultures require the addition of these reagents, or potential modifiers, at start up. Although the conceptual model on which these drawings are constructed was completed during the authors NASA/ASEE Summer Faculty Fellowship the operational prototype was not completed. Completion of the prototype is continuing at the University of Houston Clear Lake in cooperation with the NASA/JSC scientists.

A three dimensional representation of the assembled bioreactor is shown in Figure 8 to better illustrate the proposed system.

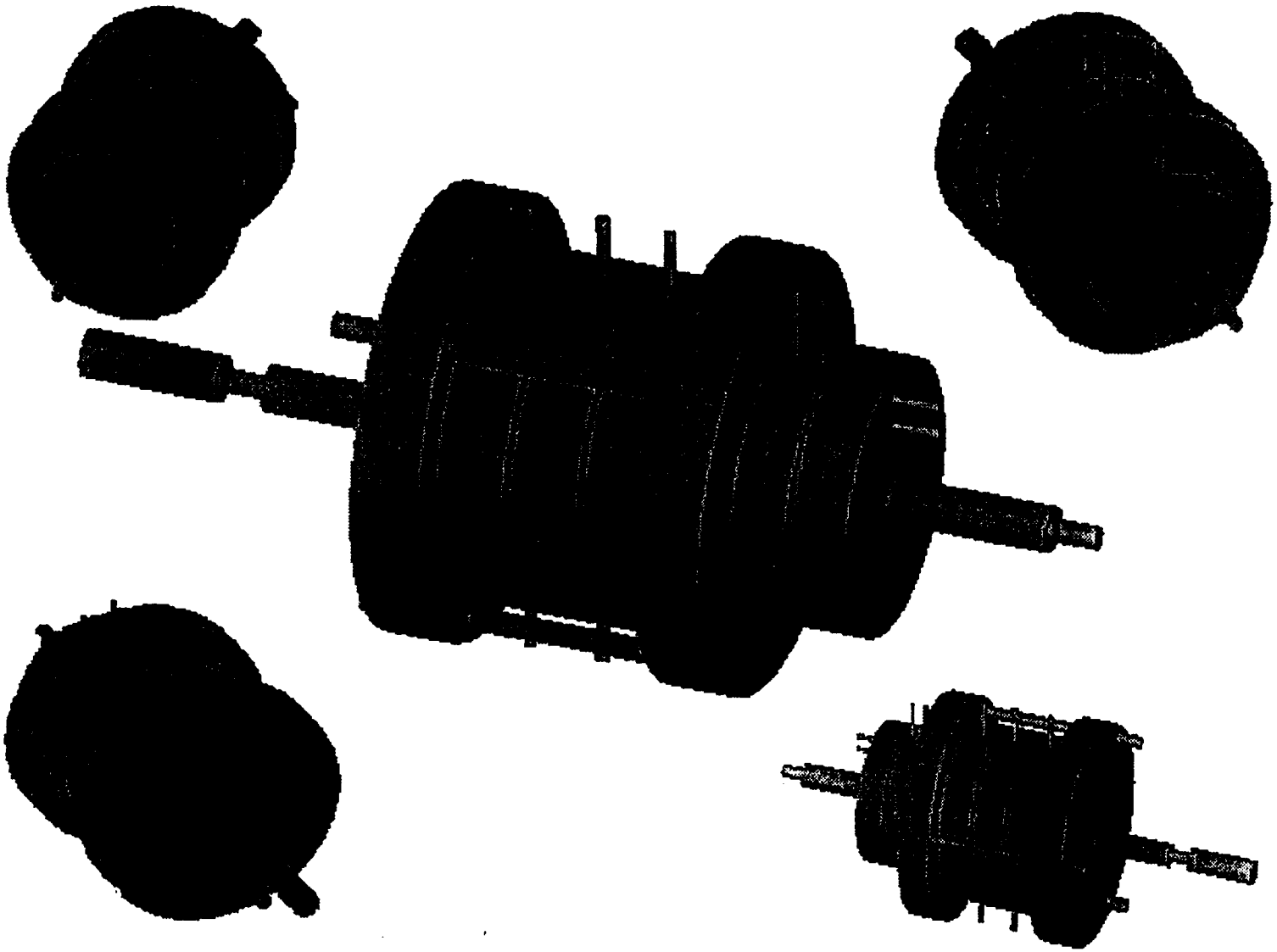


Figure 8. - Three dimensional views of operational bioreactor

REFERENCES

1. Halstead, T.W., Todd, P. and Powers, J.B., Editors: American Society for Gravitational and Space Biology. October, 1992, Volume 5, Number 2.

N94- 25369

An Evaluation of Three-Dimensional Sensors for the Extravehicular Activity Helper/Retriever

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1993

Johnson Space Center

Prepared by:	Michael Magee, Ph.D.
Academic Rank:	Professor
University & Department:	University of Wyoming Computer Science Department Laramie, Wyoming 82071-3682
NASA/JSC	
Directorate:	Engineering
Division:	Automation and Robotics
Branch:	Intelligent Systems
JSC Colleague:	Thomas W. Pendleton
Date Submitted:	August 6, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

The Extravehicular Activity Retriever/Helper (EVAHR) is a robotic device currently under development at the NASA Johnson Space Center that is designed to fetch objects or to assist in retrieving an astronaut who may have become inadvertently de-tethered. The EVAHR will be required to exhibit a high degree of intelligent autonomous operation and will base much of its reasoning upon information obtained from one or more three-dimensional sensors that it will carry and control. At the highest level of visual cognition and reasoning, the EVAHR will be required to detect objects, recognize them, and estimate their spatial orientation and location. The recognition phase and estimation of spatial pose will depend on the ability of the vision system to reliably extract geometric features of the objects such as whether the surface topologies observed are planar or curved and the spatial relationships between the component surfaces. In order to achieve these tasks, accurate sensing of the operational environment and objects in the environment will therefore be critical. This report documents the qualitative an quantitative results of empirical studies of three sensors that are capable of providing three-dimensional information to the EVAHR, but using completely different hardware approaches. The first of these devices is a phase shift laser with an effective operating range (ambiguity interval) of approximately 15 meters. The second sensor is a laser triangulation system designed to operate at much closer range and to provide higher resolution images. The third sensor is a dual camera stereo imaging system from which range images can also be obtained. The remainder of the report characterizes the strengths and weaknesses of each of these systems relative to quality of data extracted and how different object characteristics affect sensor operation.

PERCEPTRON LASER RANGE SCANNER

Operational Characteristics

The Perceptron Laser Range Scanner measures distances based on the phase shift of a modulated signal carried on an infrared laser beam. The range values returned by the scanner are represented by 12 bit integers that span a single ambiguity interval of approximately 15.2 meters. This means that a difference of one range unit (out of 4096) represents a distance change of about 4 mm. The scanner is able to produce a dense range image by employing a rotating mirror whose rotation axis can be tilted. The scanner simultaneously provides two separate range and reflectance (intensity) images that are fully registered.

Observed Characteristics of the Perceptron Laser Scanner

The quality of the range data provided by the scanner is affected by several factors which generally relate to the composition of the surface material, its reflectivity characteristics, its geometry, and the orientation of surface normals relative to the scanner itself. Without question, the most influential among these factors was the reflectivity of the surface material. For extreme cases in which a scanned region was made of a highly specularly reflective material, no reliable range estimate would be expected since the laser beam would be reflected away from the the sensor. This expectation was verified in the case of the Perceptron.

For less extreme cases involving diffuse reflective surfaces, however, the quality of the data was highly dependent on the albedo of the surface. These dependencies can best be illustrated by examining the quality of the range images acquired by scanning black and white planar surfaces (sheets of paper) that were oriented perpendicular to the optical axis of the scanner. As a measure of data stability, the local standard deviation (sigma) for range values was computed within a row. This local standard deviation was based on the center range value and the nearest 8 neighbors within the row. It was observed that the local sigma varied by as much as 3 range units. For such cases, in excess of 99% of the range samples could be expected to fall within 3 sigma (± 9 range units) of the mean value. For the test case under discussion, this translates into a local variation of approximately ±33 mm over a distance of 8 mm. For the black surface, the quality of the data was significantly worse. Local standard deviations as high as 9 range

values were observed meaning that a 3 sigma test would include range values as far as ± 100 mm over this limited region of a scan line. The local standard deviations for reflectances varied up to 30 units for the white surface and up to 8 units for the black surface.

The implications of these observed local variations are very important when designing algorithms that attempt to segment the image into component regions such as planes and curved surfaces. For example, the magnitude of the local variations in range values makes it extremely difficult to segment planar surfaces based on a local geometric constraint such as surface normal consistency. Furthermore, even on white objects, it is difficult to recognize the curvature of objects smaller than 100 mm since the magnitude of local range variation is large relative to surface size. If the data is smoothed by a classical filtering mechanism, finer details that are necessary to recognize an object and/or estimate its pose may be lost. Hence, algorithms that depend on local geometry are less likely to succeed than those that take a more global approach to object analysis. The results of both local and global algorithms that were developed are presented in the next section.

Finding Planes, Recognizing Objects and Estimating their Spatial Poses

The local instability of range values observed for the laser scanner makes scene segmentation using locally computed surface normals difficult unless the range values are smoothed using a reasonably large filter. Applying such a filter, of course, results in a loss of scene detail but does make it possible to find planes that are large relative to the size of the filter.

A better approach to finding surfaces was to grow them based on local range and reflectance difference constraints. It was determined that after applying a 7X7 mean filter, planes that were not highly oblique to the sensor axis could be successfully grown by adding to regions neighboring image elements whose smoothed reflectance and range values did not differ by more than 40 and 1.5, respectively. This provided the basis by which planar regions could be segmented and used as the basis for recognizing one of the Orbital Replacement Units (ORU) and estimating its pose using the normals of three orthogonal surfaces and a point on the handle. This feature matching method employed was based on the computed surface areas and is therefore subject to be sensitive to occlusion. Once the surfaces have been grown, however, other features could be computed such

as vertices from the intersections of planes. A large number of such features (more than four) would provide the basis for feature matching and pose estimation since there would be an overdetermined set of model/image point pairs for computing pose.

The fundamental problem with the above approach is that region growing based on propagating local constraints of reflectance and range also permits curved surfaces to be similarly grown. Attempts to differentiate between a 200 square centimeter planar surface and a 10 centimeter diameter sphere using a postprocessing step which measured the RMS error between the fitted plane and the data did not prove successful since the magnitude of range data variation was large relative to the curvature observed.

THE TECHNICAL ARTS WHITE SCANNER

White Scanner Operational Characteristics

The Technical Arts White Scanner 100 produces a dense three-dimensional range image by triangulating rays from the optical center of a camera through the image plane as they intersect a plane of laser light falling on objects in a scene. The three-dimensional coordinates thus produced are 16 bit values measured in increments of 0.001" (0.025 mm). Because the system relies on triangulating image rays through a plane of laser light, the system produces images that necessarily have missing data. This is because certain regions of the scene upon which the laser plane may be projected cannot be seen by the camera and certain regions visible to the camera may not be reached by the light plane. These missing data values are, however, flagged for the system user as invalid coordinates using a status byte.

Observed Characteristics of the White Scanner

Because the images analyzed were contained in an archive of scanned images, the extensive testing on various types of surfaces with differing albedos was not possible. However, for those images that were examined, the quality of three-dimensional data was exceptionally good. Specifically, local standard deviations rarely exceeded 1 mm. Such good resolution is, of course, to be expected since the system measures in thousandths of an

inch increments. This high quality of range measurements made it possible to find most (but not all) planes in scenes even without any data smoothing, and all planes could be found by using relatively small (3X3) smoothing filters.

The most significant observable difference between the data provided by the White Scanner and the Perceptron was noted when curved objects were scanned. Curvature was fully evident in the White Scanner images containing spheres, and their local surface normals clearly indicated vectors that converged toward the center of each sphere. This permitted the reliable use of surface normals to clearly distinguish between curved and planar surfaces.

Teleos Prism 3 Operational Characteristics

The Prism 3 Stereo Vision System uses two cameras to determine the distance to corresponding regions of interest in the left and right cameras. These regions of interest are processed in each image by applying a Laplacian-of-Gaussian operator and shifting (horizontally and vertically) the result of this operation in one image such that correspondence is established. There are two methods by which this window may be shifted. In the first case, the vergence angle between the two cameras may be adjusted, thus altering the fields of view for the cameras. The second method is to compute a stereo disparity by shifting the window (in software) until a match is obtained. For producing a dense range image, the second method is preferred because no mechanical repositioning of the camera system is required.

THE TELEOS PRISM 3 STEREO VISION SYSTEM

Observed Characteristics of the Prism 3

The Prism 3 produces the best quality of measurements when highly visually textured objects are being observed. For test purposes, two such objects, a cylinder and a planar surface, were covered with a white material upon which there was a black speckled pattern. The cylinder and planar surface were observed under varying conditions relative to each other with the following results.

When the cylinder was placed in contact with the plane at a distance of about 2 meters, the curvature of the cylinder was clearly evident, although it blended into the background plane in a manner resembling a Gaussian curve. In this configuration, the camera was both focused and verged on the cylinder.

With the cylinder remaining in the same position, the plane was moved further away from the Prism 3. As it moved further away, significant anomalies in the data developed, especially near the (jump) boundaries of the cylinder. With the background plane at a very large distance (effectively removed from the scene), the cylinder lost a significant amount of its observed curvature and the spikes near the edge of the cylinder became very large.

Although significant deterioration in data quality was noticed as the separation between the plane and cylinder increased, it was found to be possible to use internal correlation measures computed by the Prism 3 as the basis for confidences that could distinguish between good and bad data. The confidence measures were based on correlation peak strength, distinctiveness, and uniqueness. Examining images in the context of these measures revealed that although spurious measurements existed at the object boundaries, their confidences were low relative to measurements not near boundaries.

In order to obtain an estimate of RMS errors for the raw data itself, range values for the solitary plane were used. It was found that at distances of approximately 2 meters the standard deviation for range values was less than 0.8 cm.

It should be noted in all the tests for the Prism 3 that accurate focus and vergence are critical in order to obtain the best results. Hence, for the tests involving two objects (as in the cases involving the cylinder and plane), unless the lenses used have a large depth-of-focus, the range values for one object will necessarily deteriorate. Similarly, if the system is verged on the forward object, then vergence on the rear object becomes worse as it moves further away.

N94- 25370

DEVELOPMENT OF PROGRAMMABLE ARTIFICIAL NEURAL NETWORKS

Final Report

NASA/ASEE Summer Faculty Fellowship Program–1993

Johnson Space Center

Prepared By:	Andrew J. Meade, Jr., Ph.D
Academic Rank:	Assistant Professor
University & Department	Rice University Department of Mechanical Engineering and Materials Science Houston, Texas 77251-1892
NASA/JSC	
Directorate:	Information Systems
Division:	Information Technology
Branch:	Software Technology
JSC Colleague:	Robert O. Shelton, Ph.D
Date Submitted:	September 30, 1993
Contract Number:	NGT-44-001-800
Approved By:	Robert O. Shelton
Date Approved:	10/14/93

DEVELOPMENT OF PROGRAMMABLE ARTIFICIAL NEURAL NETWORKS

Final Report

NASA/ASEE Summer Faculty Fellowship Program–1993

Johnson Space Center

Prepared By:	Andrew J. Meade, Jr., Ph.D
Academic Rank:	Assistant Professor
University & Department	Rice University Department of Mechanical Engineering and Materials Science Houston, Texas 77251-1892
NASA/JSC	
Directorate:	Information Systems
Division:	Information Technology
Branch:	Software Technology
JSC Colleague:	Robert O. Shelton, Ph.D
Date Submitted:	September 30, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

Conventionally programmed digital computers can process numbers with great speed and precision, but do not easily recognize patterns or imprecise or contradictory data. Instead of being programmed in the conventional sense, artificial neural networks are capable of self-learning through exposure to repeated examples. However, the training of an ANN can be a time consuming and unpredictable process.

A general method is being developed by the author to mate the adaptability of the ANN with the speed and precision of the digital computer. This method has been successful in building feedforward networks that can approximate functions and their partial derivatives from examples in a single iteration. The general method also allows the formation of feedforward networks that can approximate the solution to nonlinear ordinary and partial differential equations to desired accuracy without the need of examples. It is believed that continued research will produce artificial neural networks that can be used with confidence in practical scientific computing and engineering applications.

INTRODUCTION

Neural networks have proven to be versatile tools for accomplishing what could be termed higher order tasks such as pattern recognition, classification, and visual processing. However, conventional wisdom has held that networks are unsuited for use in more purely computational tasks, such as mathematical modelling and physical analysis of engineering systems. Certainly the biological underpinnings of the neural network concept suggest that networks would perform best at tasks at which biological systems excel, and worse or not at all at other tasks.

Contrary to popular opinion the author believes that continued research into the approximation capabilities of networks will enable the neural network paradigm, with all of its advantages in behavior and adaptability, to be mated to the more purely computational paradigms of mathematically oriented scientific programming and analysis. Additionally, it is felt that the thorough investigation of network approximation capabilities will benefit the network field and connectionism in general.

In a field as conceptually difficult as the study of artificial neural networks, it is best to start investigation with supervised learning, test the established premises, and alter them to circumvent pitfalls in implimentation.

FUNCTION APPROXIMATION

Learning as Function Approximation

Central to the author's research approach is the view that supervised learning in artificial neural networks is equivalent to the problem of approximating a multivariate function and that learning should be able to be explained by approximation theory. Approximation theory deals with the problem of approximating or interpolating a multivariate function. This approach has been considered by other researchers in the field of ANNs [1]-[4]. However, the author extends this assumption of function approximation by assuming that ANNs can model discontinuous multivariate functions and should be at least as accurate and numerically efficient as existing computational techniques used in science and engineering. Also, ANN behavior and adaptation difficulties, from supervised learning to machine vision, should be amenable to the standard error analysis techniques used in numerical analysis [5].

Function Approximation In Engineering

There are three classes of tools used in science and engineering for the analysis of systems:

1. **analytical methods**, which include the formation of equations that model the behavior of systems and the analytic solution of those equations.

2. **computational methods**, which involve the simulation of system behavior by the numerical solution of the governing equations.

3. **experiments**, which involve the investigation of physical phenomena and the gathering of data to validate analytical models and numerical simulations.

In a general sense analytical and computational methods and experiments can be considered to be forms of function approximation. The governing equations derived from analytical methods are a compact representation of the functions that model some particular phenomena observed in experiments. Computational techniques are used to approximate the function or functions that satisfy the governing equations. The graphs and tables made from experiments are representations of the functions that underlie observed physical phenomena.

Computational Methods

In the wake of the computer revolution in scientific applications, a large number of computational techniques have emerged. Also, particular methods have assumed prominent positions in certain areas of application. For example, finite element methods are used almost exclusively for solving structural problems; spectral methods are becoming the preferred approach to global atmospheric modelling; and the use of finite difference methods is nearly universal in simulating fluid and thermal systems.

Each computational method has its own set of advantages and disadvantages depending on the characteristic of the application. These popular and apparently unrelated techniques are firmly entrenched in computer codes used every day by practicing scientists and engineers. Often the formal numerical training provided the scientist and engineer reinforces the divisions between the various computational methods available. However, Fletcher [6] has demonstrated that each of these numerical methods are in fact particular aspects of a more general approach known as the **method of weighted residuals** [7].

PROGRAMMABLE ARTIFICIAL NEURAL NETWORKS

It is the objective of the author's research program to demonstrate that artificial neural network behavior from supervised learning to machine vision can be derived from the method of weighted residuals. This would link ANNs with the relatively mature and established field of computational mechanics, extend ANN capabilities, and help in transforming ANN applications from an art to a science. This may also advance research in our understanding of biological neural systems.

If we are to assume that ANNs are as valid as established computational techniques, then ANNs should be evaluated in the same manner as are computational techniques.

The first step in evaluating the capabilities of a new numerical method, is to apply it to the solution of algebraic and ordinary and partial differential equations of known behavior. This same approach can be used for ANNs since the solution of algebraic and differential equations can be viewed as the approximation of a function that must satisfy the equation in question subjected to boundary and/or initial conditions.

Applying an ANN to the solution of an algebraic or differential equation effectively uncouples the influences of the quality of data samples, network architecture, and transfer functions from the network approximation performance. The solution of equations also allows us to study the influence of constraining the connection weights. The most immediate benefit in this approach would be the construction of networks that can approximate the solution to desired equations without the need for examples. This would be of value in

engineering applications since considerable effort may be saved if the equations governing a physical process can be directly incorporated into the neural network architecture without the need of examples, thereby shortening or even eliminating the learning phase.

This approach may also lead to the construction of network training routines that are faster and more accurate than those presently in use [8], [9]. In addition, progress made in this network programming approach should provide the research community insight into the working of networks for associative memory, classification, and machine vision applications.

Approach

The MWR approach has been taken by the author using the hard limit [10] as the transfer function. Interesting results have been produced and are presented to demonstrate the validity of the approach.

It can be argued that the supervised training of a feedforward network is a problem in function approximation using unconstrained optimization. In this sense, the task of the optimization scheme is to find the proper combination of connection weights between the processing elements, operating with specific transfer functions, so that the network minimizes the error between the network output and the desired output. Therefore, the training of a network possesses all of the problems one associates with unconstrained optimization such as avoiding local minima in search of the global minimum.

The most obvious remedy to this problem is to constrain the optimization while preserving the approximation capability of the network. Our task then is to form constraints between the weights so that the values of the weights may be determined with computational efficiency.

Constraint of Weights Between the Input and Hidden Layer

A univariate function $u(x)$ can be represented by a feedforward network (Fig. 1) with a single hidden layer, and a single input and output node using a linear transfer function, as follows:

$$u(x) = \sum_{i=1}^{N} \left(\Phi_i\left(\xi_i\right) + s_i\right) w_i, \quad \xi_i = \alpha_i x + \theta_i \quad \text{for} \quad i = 1, ..., N \tag{1}$$

Each hidden processing element is indexed by the subscript i, where N is the total number of hidden processing elements. The variable x is the value of the input and Φ_i represents the nonlinear transfer function for the i^{th} hidden processing element. The coefficients w_i, α_i and θ_i are the values of the connection weights between the hidden and output layers, the input and hidden layers, and the bias node and hidden layer, respectively. The role of the remaining set of coefficients, s_i, will be explained shortly. Equation (1) indicates that we must determine the value of $4N$ coefficients to approximate a function with the feedforward architecture. Our objective in constraining the weights is to decrease the number of unknown coefficients.

The formulation of Eq. (1) is similar to that given by Cybenko [11]. However, ξ_i will be modified in a manner similar to the radial basis technique. Notice that to provide extended dynamic range, this formulation assumes that the input and output processing elements

use linear transfer functions. However, this formulation also allows the use of nonlinear transfer functions in the output node.

We will constrain the weights θ_i in the following manner. Discretize the domain (Ω) of the input variable into $N-1$ intervals. Each interval is bracketed by the values $\hat{x}_i$ and $\hat{x}_{i+1}$ where $\hat{x}_i < \hat{x}_{i+1}$ and $i = 1, ..., N$. The value of N is equal to the total number of hidden processing elements. We will use the following equation to constrain θ_i:

$$\theta_i = -\alpha_i \hat{x}_i \quad \text{for} \quad i = 1, ..., N$$

so that

$$\xi_i = \alpha_i (x - \hat{x}_i) \quad \text{for} \quad i = 1, ..., N.$$

Notice that by constraining each bias weight (θ_i) in this manner, $\xi_i = 0$ for the i^{th} processing element when $x = \hat{x}_i$. The variables $\hat{x}_i$ are similar to the "centers" used in the radial basis function literature [12],[13].

For our analysis we will use a piecewise continuous polynomial approximation to the hyperbolic tangent (Eq. (2)) known as the hard limit which is illustrated in Fig. 2.

$$\Phi_i(\xi_i) = \xi_i - 1 \quad \text{where} \quad \xi_i = \frac{2(x - \hat{x}_i)}{\hat{x}_{i+1} - \hat{x}_i} \quad \text{for} \quad \hat{x}_i \leq x \leq \hat{x}_{i+1}$$

$$\Phi_i(\xi_i) = -1 \quad \text{for} \quad x \leq \hat{x}_i \quad \text{and} \quad \Phi_i(\xi_i) = +1 \quad \text{for} \quad \hat{x}_{i+1} \leq x. \tag{2}$$

Therefore, for the piecewise polynomial transfer function we can constrain the input weights α_i using Eq. (2).

$$\alpha_i = \frac{2}{\hat{x}_{i+1} - \hat{x}_i} \quad \text{for} \quad \hat{x}_i \leq x \leq \hat{x}_{i+1}$$

Notice then, that transfer functions from each processing element are distributed along the x axis in the domain of interest Ω. Each function is "centered" at the respective values of $\hat{x}_i$ (Fig. 3).

Equation (1) can now be seen as a weighted sum representation of the function $u(x)$. One can think of the transfer functions as being interpolation functions distributed along the transformed x axis. Each interpolation function, and its coefficient s_i, is multiplied by its respective weight, w_i, and summed to approximate a desired curve.

Constraint of Weights Between the Hidden and Output Layer

It can now be shown that the role of the coefficients s_i in Eq. (1) is to add or subtract constant values from the respective transfer function. This effectively moves the transfer function above or below the x axis (Fig. 4). Equation (1) may be rewritten as:

$$u(x) = \sum_{i=1}^{N} \Phi_i(\xi_i) w_i + \sigma \quad \text{where} \quad \sigma = \sum_{i=1}^{N} s_i w_i \tag{3}$$

The coefficient σ acts as the connection weight between the second bias and the output node (Fig. 1).

One final constraint we impose on the output weights in that $w_i = -w_{i+1}$. This requires that we use an even number of processing elements. The final constraint acts to convert the hard limits from global interpolation functions into local interpolation functions.

We have now constrained all of the parameters for the single input and single output feedforward network and have decreased the number of unknowns from $4N$ (α_i, θ_i, s_i, and w_i) to $N/2+1$ (w_i and σ). What remains now is to determine the output weights and the second bias weight so as to approximate not only the desired functional relationship, but also the derivatives of the function. This is done using the method of weighted residuals (MWR). More specifically, the Bubnov-Galerkin and Petrov-Galerkin methods.

ORDINARY DIFFERENTIAL EQUATIONS

Determination of Output Weights: Method of Weighted Residuals

To illustrate the method we will determine the output weights and bias (w_i and σ) needed to approximate the solution to a linear first order differential equation, using only the equation and the initial condition [14].

A feedforward network of one input and output node and a single hidden layer is constructed to approximate the solution to the following equation:

$$\frac{du}{dx} - u = 0 \tag{4}$$

with the boundary condition $u(0) = 0$. The exact solution is:

$$u_{\text{exact}} = e^x$$

Substituting Eq. (1) into Eq. (4) we have:

$$\frac{du}{dx} - u = \sum_{i=1}^{N} \frac{d\Phi_i(\xi_i)}{dx} w_i - \sum_{i=1}^{N} \Phi_i(\xi_i) w_i + \sigma = 0. \tag{5}$$

In this example we will set σ to zero. Notice that Eq. (5) is only satisfied if the nontrivial set w_i is exactly correct.

If we use random values for w_i the right hand side would be nonzero and act as a measure of the error. This error is also known as the equation residual (ϵ). We may obtain an acceptable approximation if we force the weighted residual to zero over the domain of interest, Ω.

$$\int_\Omega f_k(x)(\epsilon)\,dx = 0 = \int_\Omega f_k(x) \left(\sum_{i=1}^{N} \frac{d\Phi_i(\xi_i)}{dx} w_i - \Phi_i(\xi_i) w_i \right) dx \tag{6}$$

where $k = 1, \ldots, N$ and $f_k(x)$ is referred to as the weighting function or test function. The approach shown by Eq. (6) is known as the method of weighted residuals.

Since linearly independent relationships are needed to solve for the coefficients w_i, it is clear that f_k must be a set of linearly independent functions. The choice of the weighting functions correspond to different solution techniques such as the subdomain, collocation, least square, and the Galerkin methods [6]. For this investigation we will use a modification of the Galerkin approach or more specifically the Bubnov-Galerkin method [15]. This approach requires that f_k be chosen from the same family of functions as the transfer functions, that is,

$$f_k(x) = \Phi_k(\xi_k) \quad \text{for} \quad k = 1, ..., \ N. \tag{7}$$

Except for the change in the index from i to k, $\Phi_k(\xi_k)$ is described by the hat function.

So Eq. (6) may be rewritten as

$$\sum_{i=1}^{N} \int_{\hat{x}_1}^{\hat{x}_N} \Phi_k(\xi_k) \frac{d\Phi_i(\xi_i)}{dx} \, dx \ w_i - \sum_{i=1}^{N} \int_{\hat{x}_1}^{\hat{x}_N} \Phi_k(\xi_k)\Phi_i(\xi_i) \, dx \ w_i = 0 \quad \text{for} \quad k = 1, ..., N. \tag{8}$$

This forms the linear algebraic system of equations

$$\sum_{i=1}^{N} A_{ki} w_i = g_k \quad \text{for} \quad k = 1, ..., N. \tag{9}$$

In its present form A_{ki} of Eq. (9) is singular and must be modified by the initial condition. The initial condition may be written as

$$\sum_{i=1}^{N} \Phi_i(\hat{x}_1) \, w_i = g_1 = 1. \tag{10}$$

The weights w_i can then be evaluated directly from the solution of Eq. (9). Notice that the initial condition could have been partially satisfied by w_i and the value of σ would have been determined to satisfy the remainder (i.e. $g_1 = 0.5$ and $\sigma = 0.5$)

Comparison of the network approximation with the exact solution of Eq. (4) is shown in Fig. 5 for forty two hidden processing elements (twenty one output weights).

Example: Third Order Nonlinear Ordinary Differential Equation

As a further demonstration of the approximation capability of the network, a feedforward network of one input and output node, and a single hidden layer, was constructed to approximate the solution to the nonlinear third order ordinary differential equation known as the Blasius equation [17]. The Blasius equation is used to describe the steady and laminar two-dimensional flow of a viscous Newtonian fluid about a flat plate:

$$\frac{d^3 f}{d\eta^3} + f \frac{d^2 f}{d\eta^2} = 0 \tag{11}$$

with boundary conditions

$$f(0) = \frac{df}{d\eta}(0) = 0, \ \frac{df}{d\eta}(\eta \to \infty) = 1$$

The network approximation (Fig. 6), using fifty one output weights, is compared against a fifth order Runge-Kutta solver (finite difference) of variable step size that satisfied the boundary conditions with an absolute error value of 10^{-6}. Figure 7 illustrates the rate of convergence of the network using the $L2$ norm of the error and the interval spacing h.

PARTIAL DIFFERENTIAL EQUATIONS

The author has successfully programmed Higher Order Networks, also known as Sigma-Pi networks, to approximate the solution of a partial differential equation [16].

Example: Linear Elliptic Partial Differential Equation

A Sigma-Pi network of two inputs, one output node, and a single hidden layer has been constructed to approximate the solution to the linear elliptic equation that models fully developed steady flow of viscous Newtonian fluid through a duct of square cross-section:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + 1.0 = 0.0 \tag{12}$$

where u is the nondimensionalized velocity of the flow along the duct.

The domain (Ω) of the problem is: $-1.0 \leq x \leq 1.0$, $-1.0 \leq y \leq 1.0$

Boundary conditions: $u = 0$ along the perimeter Γ

Figures 8 and 9 show the surfaces made by the exact solution and the network solution of Eq. (12) for two thousand five hundred output weights. Figure 10 illustrates the rate of convergence of the network using the root mean square of the error and the interval spacing h.

REFERENCES

[1] Barron, A.R. and Barron, R.L., "Statistical learning networks: a unifying view," *Symposium on the Interface: Statistics and Computing Science,* Reston, Virginia, April 1988.

[2] Omohundro, S., "Efficient algorithms with neural network behaviour," *Complex Systems*, **1**, 237, 1987.

[3] Poggio, T. and Girosi, F., "A Theory for Networks for Approximation and Learning," *Artificial Intelligence Laboratory, Massachusetts Institute of Technology*, A.I. Memo No. 1140, July 1989.

[4] Girosi, F. and Poggio, T., "Networks for learning: a view from the theory of approximation of functions," *Proceedings of the Genoa Summer School on neural networks and their applications*, Prentice Hall, 1989.

[5] Atkinson, K.E., *An Introduction to Numerical Analysis*, John Wiley & Sons, New York, 1989.

[6] Fletcher, C.A.J., *Computational Galerkin Methods*, Springer-Verlag, New York, 1984.

[7] Finlayson, B.A., *The Method of Weighted Residuals and Variational Principles*, Academic Press, New York, 1972.

[8] Freeman, J.A. and Skapura, D.A., *Neural Networks: Algorithms, Applications, and Programming Techniques*, Addison-Wesley Publishing, 1991.

[9] Hecht-Nielsen, R., *Neurocomputing*, Addison-Wesley Publishing, 1990.

[10] Maren, A., Harston, C. and Pap, R., *Handbook of Neural Computing Applications*, 48, Addison-Wesley Publishing, 1990.

[11] Cybenko, G. "Approximation by Superposition of a Sigmoidal Function," Math. Control Signals Systems, **2**, 303-314, (1989).

[12] Elanayar, S. and Shin, Y. C., "Approximation Capabilities of Radial Basis Function Neural Networks," *Intelligent Engineering Systems Through Artificial Neural Networks,* Vol. 2, pp. 291-298, ASME Press, New York, 1992.

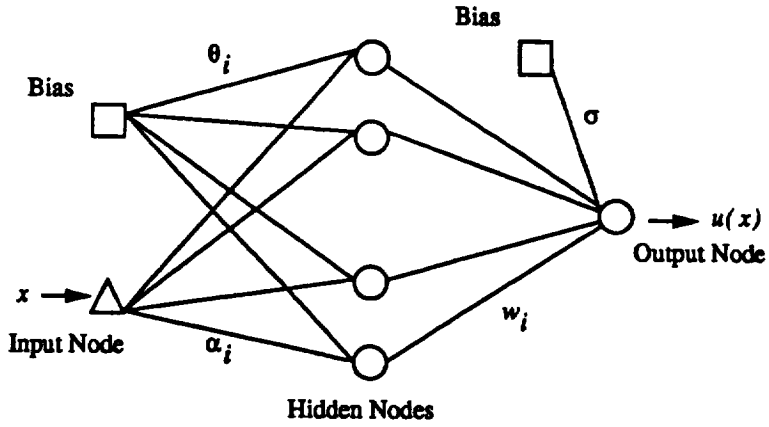
[13] Park, J. and Sandburg, I. W., "Universal Approximation Using Radial Basis Function Networks," *Neural Computation*, **3**, 246-257, 1991.

[14] "Solution of Ordinary Differential Equations by Programmable Feedforward Neural Networks," To be submitted, 1993.

[15] Mikhlin, S. G., *Variational Methods in Mathematical Physics*, Pergamon, Oxford, 1964.

[16] "Solution of Elliptic, Parabolic and Hyperbolic Partial Differential Equations by Programmable Feedforward Neural Networks," To be submitted, 1993.

[17] White, Frank M., *Viscous Fluid Flow*, McGraw-Hill, New York, 1974.



θ_i : Bias weight for the i^{th} hidden node.

α_i : Input weight for the i^{th} hidden node.

w_i : Output weight for the i^{th} hidden node.

σ : Bias weight for the output node.

Figure 1: Feedforward Network Architecture

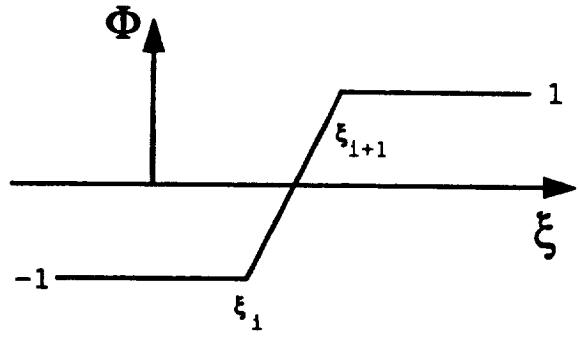


Figure 2: Piecewise Polynomial Transfer Function

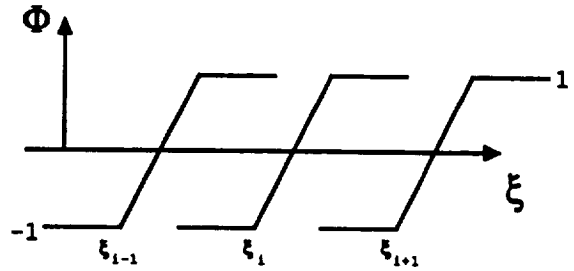


Figure 3: Distribution of Transfer Functions

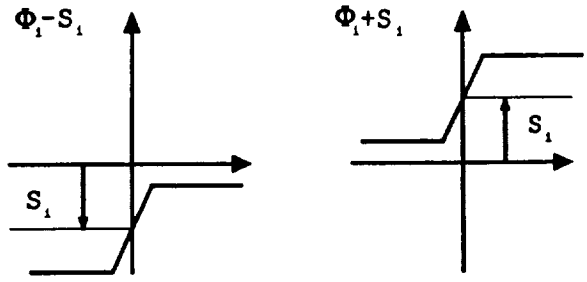


Figure 4: Transfer Function Offset

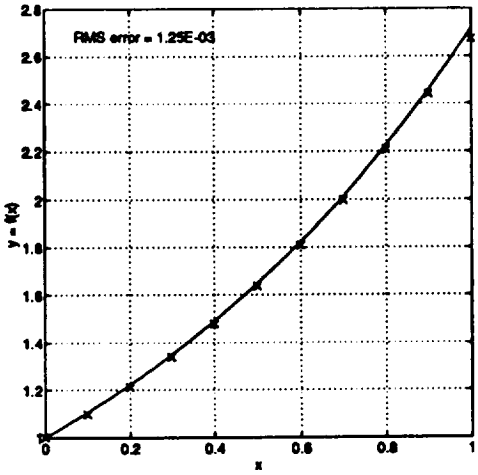


Figure 5: Comparison of Exact Solution and Network Approximation of Eq. (4) for 21 output weights

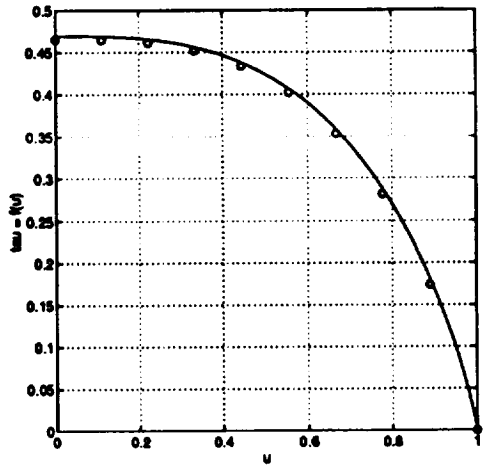


Figure 6: Comparison of Runge-Kutta Solution and Network Approximation of Eq. (11), for 51 output weights

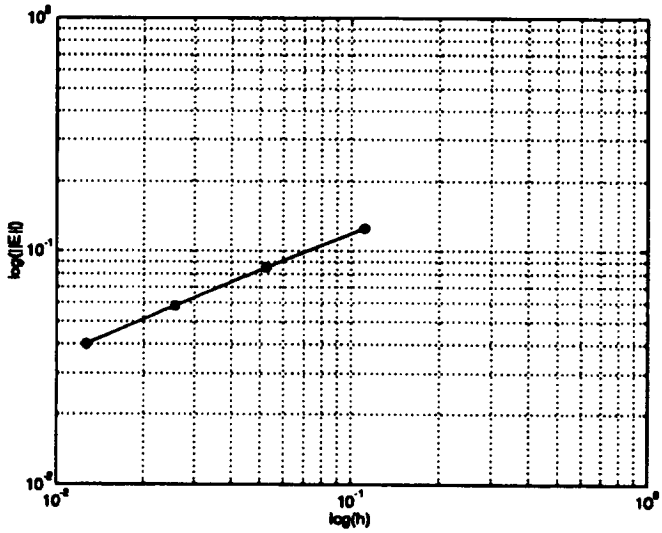


Figure 7: Convergence Rate of Network for Eq. (11)

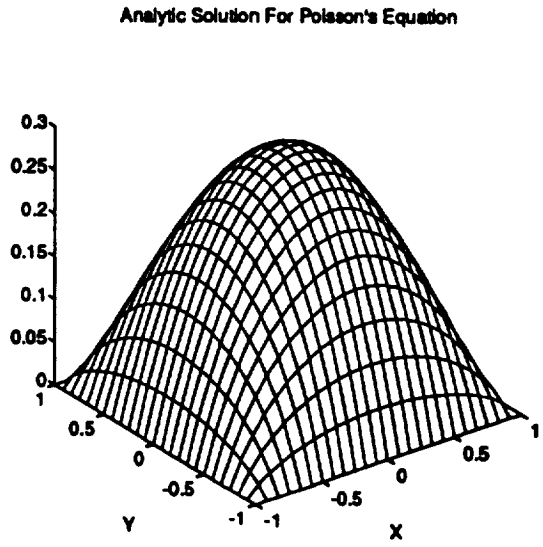


Figure 8: Analytic Solution of Eq. (12)

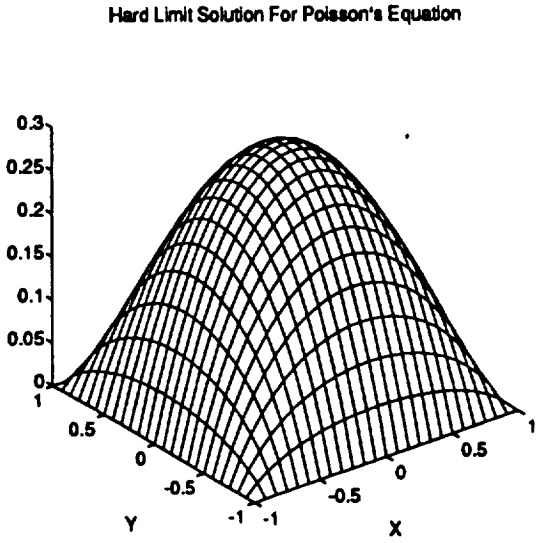


Figure 9: Network Solution of Eq. (12) for 2500 output weights

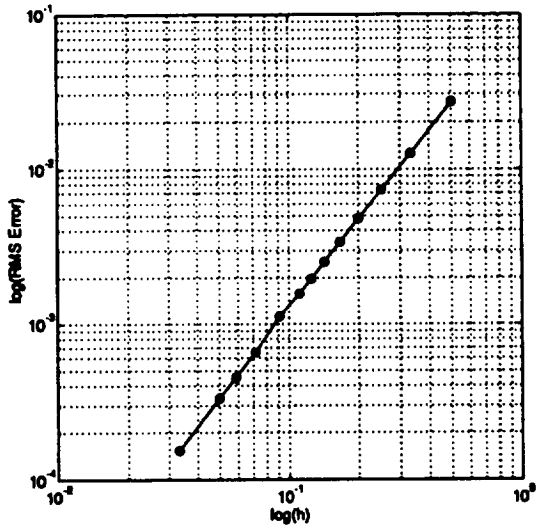


Figure 10: Convergence Rate of Network for Eq. (12)

N94- 25371

ROBOTICS IN A CONTROLLED, ECOLOGICAL LIFE SUPPORT SYSTEM

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Gaines E. Miles, Ph.D	Kimberly J. Krom, Student
Academic Rank:	Professor	Agricultural Systems Management
University & Department:	Purdue University Department of Agricultural Engineering West Lafayette, Indiana 47907-1146	
NASA/JSC		
Directorate:	Engineering	
Division:	Crew and Thermal Systems	
Branch:	Life Support Systems	
JSC Colleague:	Daniel J. Barta, PhD	
Date Submitted:	August 5, 1993	
Contract Number:	NGT-44-001-800	

ROBOTICS IN A CONTROLLED, ECOLOGICAL LIFE SUPPORT SYSTEM

ABSTRACT

Controlled, Ecological Life Support Systems (CELSS) that utilize plants to provide food, water and oxygen could consume considerable amounts of labor unless crop production, recovery and processing are automated. Robotic manipulators equipped with special end-effectors and programmed to perform the sensing and materials handling tasks would minimize the amount of astronaut labor required.

The Human Rated Test Facility (HRTF) planned for Johnson Space Center could discover and demonstrate techniques of crop production which can be reliably integrated with machinery to minimize labor requirements. Before the physical components (shelves, lighting fixtures, etc.) can be selected, a systems analysis must be performed to determine which alternative processes should be followed and how the materials handling tasks should be automated.

Given that the current procedures used to grow crops in a CELSS may not be the best methods to automate, then what are the alternatives? How may plants be grown, harvested, processed for food, and the inedible components recycled? What commercial technologies current exist? What research efforts are underway to develop new technologies which might satisfy the need for automation in a CELSS? The answers to these questions should prove enlightening and provide some of the information necessary to perform the systems analysis.

The planting, culturing, gathering, threshing and separation, food processing, and recovery of inedible portions of wheat were studied. The basic biological and materials handling processes of each task are defined and discussed. Current practices at Johnson Space Center and other NASA centers are described and compared to common production practices in the plant production industry. Technologies currently being researched which might be applicable are identified and illustrated. Finally, based on this knowledge, several scenarios are proposed for automating the tasks for wheat.

ROBOTICS IN A CONTROLLED, ECOLOGICAL LIFE SUPPORT SYSTEM

INTRODUCTION

Controlled, Ecological Life Support Systems (CELSS)

The purpose of a CELSS is to provide food and to replenish supplies of oxygen and water for long-term voyages in space. The basic physiological processes of plant growth (photosynthesis, nutrient uptake, translocation, transpiration, and respiration) combine carbon dioxide, minerals and water to form food and inedible plant materials which must be recycled. Photosynthesis removes carbon dioxide exhaled by the astronauts into the atmosphere and supplies oxygen for breathing. Transpiration adds odorless moisture to the air, where the water can be condensed and purified for consumption.

The successful culture of higher-order plants requires the careful management of complex biological processes. A CELSS must have seeds (or other propagules), a source of mineral nutrients, water, equipment, radiant energy (light), and labor or machinery to move materials from one process to another. Artificial intelligence is necessary to monitor the state of the CELSS and diagnose and correct identified problems. This report describes and proposes technologies useful for automating CELSS.

Robotics

A robot is a multifunctional manipulator which can be reprogrammed to perform a variety of materials handling tasks. Robotics includes the sensors and artificial intelligence programs which enable the machine to perceive changes in its surrounding environment and to respond appropriately. Computer programs collect and process data, and characterize the resulting information to create the knowledge base for determining what actions the robot should take. The interface between plants and the robot is the gripper, or end-effector which is probably unique to each task and to each crop.

Sensors

Electrical sensors provide sight, smell and touch which enables the robot to perform the desired tasks, even in unpredictable environments. Examples include image processing to locate fruit, detect and grade seedlings, monitor plant health and to determine fruit ripeness.

The location of fruit or other plant parts to be harvested can be determined by machine vision and image processing. A machine vision system typically consists of a camera with a CCD sensor chip, optical lenses to provide the desired field of view, and a frame grabber or digitizer interface to a computer. Image processing programs convert the raw image data from the frame grabber into knowledge about the scene. A circular Hough transform has been used to locate tomatoes (Whittaker, *et al.*, 1987). A projected plane of laser light was used by Benady and Miles (1992) reduce the amount of image processing required to locate muskmelons in the time required by a robotic harvester. Simonton (1990) and Guyer, *et al.* (1986) developed machine vision and image processing algorithms to measure plant features.

Hetzroni, *et al.* (1992) used a machine vision and image processing system to monitor the nutrition and health of lettuce grown in a controlled environment chamber. Neural networks were used to classify picture elements (pixels) into normal or nutrient-deficient classes. Miles (1989; 1991) showed that image processing could detect patterns in wheat leaves caused by nitrogen, iron and/or potassium deficiencies.

The ROTRAN® 2000[1] robotic transplanter for bedding plants uses an integrated machine vision and image processing system to check for seedlings and to direct the robot to correct for misses (Beam, *et al*, 1991). This is one of the few commercially available machines which has the versatility required of a CELSS.

Benady, *et al.*, 1992 utilized an electronic sensor for ethylene to determine ripeness of cantaloupes. The aromatic volatile gases emitted naturally from climacteric fruit during ripening are detectable by the small, hand-held SnO_2 sensor. The ability to accurately measure crop ripeness is necessary for selective harvesting, and for automating a CELSS.

Manipulators and End-effectors

Robot manipulators may be Cartesian, revolute, or hybrid combinations depending on the physical location of the servomechanisms. In Cartesian style robots, the actuators are positioned so that each axis provides linear motion for a carriage that carries the next axis, or the end-effector. Curvilinear motion is provided by coordinating the relative motion between each axis. In revolute robots, the manipulator consists of a base and arm sections with servos at each joint to provide rotary motion. By coordinating the motions between each servo, linear motions of the end-effector can be achieved. Cartesian-axes and revolute-joints may be combined to produce a wide variety of robot types.

End-effectors, or grippers provide the mechanisms to grasp objects. Many end-effectors also have axes or joints, which permit objects to be positioned or oriented independently of translocation by the manipulator. The physical design and size of grippers depends on the objects to be grasped, and may be unique for each task. Simonton (1991) has developed end-effectors for manipulating plant materials.

Because the robot(s) will be required to perform a multitude of tasks, the end-effectors must be automatically changeable, without human assistance. This capability is not normally found on industrial robots and may require considerable development efforts for the unique set of CELSS end-effectors. The connections to the end-effector must include:

- physical support,
- electrical, pneumatic, vacuum, and/or hydraulic service, as well as
- sensor and control lines.

Transplanting

Kutz, *et al.* 1987 demonstrated robotic transplanting of bedding plants from a seedling flat to a grow flat (Miles and Kutz, 1991). The gripper was two flat pieces of spring steel that open and closed pneumatically around the seedling plug. The Puma robot inserted the gripper into each cell by following an "L" shaped approach path which kept

[1]ROTRAN 2000 is a registered trademark of Robotic Solutions, Inc., 1291-G Cumberland Avenue, West Lafayette, IN 47906.

the seedlings from being bent and broken during the downward motion. The seedlings were inserted in a previously dibbled hole in the grow flat soil mix, and released by opening the grippers.

Beam, *et al.*, 1991, developed a Cartesian type of robotic transplanter which uses a similar gripper, but employs different technologies for automating and coordinating motions of the twin gantries and the conveyors which position the flats. Pneumatics are used for positioning the gripper vertically and for opening the fingers, while stepper-motors are used to position the gantry carriages laterally, and to position the flats on the conveyors. Because all the motions are controlled by a microcomputer, ROTRAN® 2000 is capable of transplanting from a wide variety of flat types and sizes, at a rate of approximately 2000 seedlings per hour.

Melon Harvesting

Selective harvesting of cantaloupes is being accomplished by a Cartesian robot designed by engineers at Purdue University and the Agricultural Engineering Research Institute, Agricultural Research Organization (Volcani Center), Bet Dagan, Israel (Edan and Miles, 1991; Edan, *et al.*, 1991; Benady, *et al.*, 1991). The 3 axes position the ring grippers over the fruit, and match the ground speed of the vehicle while the gripper descends and grasps each melon. After the rings close around the melon and the manipulator lifts it a short distance, a swinging knife trims the vine. Because the manipulator picks each fruit individually and the machine vision and image processing system and the "sniffer" sensors detect fruit ripeness, multiple and selective harvests are possible with this robot.

Animated Simulation/Systems Engineering

Because of the versatility of programmable machines and the complexity of tasks, a myriad of scenarios are possible when automating crop production. When harvesting a row, or tree, are differences in performance from one test to another due to changes made in the machine, or differences in physical properties of the crop? In agriculture, it's not uncommon for threshing efficiency to change by 30% or more due to changes in the physical properties of wheat, with no changes in combine settings. Such uncertainties could lead to endless experimental research. A preferred approach would be to develop 3-D, animated models of robots which could be used to simulate materials handling and quantify responses for changes in environmental conditions and the machine design. This approach would permit the affects of changes in design of the machine to be clearly separated from crop parameters. In addition, solutions are much quicker to obtain, and answers can be obtained at any time, not just when the crop is ready.

The animated simulation approach clearly defines the work space and time requirements to perform a task. This information determines some critical design parameters for the robot, namely the size or length of each axis, and the sizes of servos required to move the materials along each axis in the allocated time period. With this information, the design engineer will be able to make better decisions concerning the robot hardware and software specifications.

The modeling and simulation approach does not eliminate the need for laboratory studies to validate the results. Tests must be performed to confirm that the grippers and

sensors perform as intended. But, the number of tests are greatly reduced, since the simulations will reveal infeasible solutions. Thus, simulation is a tool which is best used in conjunction with a testing program.

In automating a CELSS, there will be distinct tradeoffs among complexity of end-effectors, reliability, speed of performance, and mass for each task. Obviously, end-effectors must be light-weight to reduce launch mass, and reliable to reduce crew interdiction, but how fast must the end-effector do it's job? That depends on whether or not the robot has other jobs waiting. If the robot would otherwise be idle, its possible to use a simple, but slow end-effector, which probably would be much lighter.

Artificial Intelligence

A functioning CELSS is a complex system, and when astronauts are depending on it for life support, it is critical that each process be monitored and controlled. Failure of a single component or process could prove disastrous. To prevent catastrophic failures, artificial intelligence programs must be written to ascertain that the CELSS performance is on-course, off-course but correcting back on line, or off-course far enough to be lost and jeopardizing the mission. Numerical models, simulations, expert systems, neural networks and other aspects of artificial intelligence will be required to predict on-course or ideal conditions and to detect departures.

Crop Growth Simulation

Alfalfa has been modeled as a set of simultaneous equations which explain the rate of processes such as photosynthesis, respiration, translocation, transpiration, etc. as functions of both meteorological conditions and crop parameters (Holt, et al., 1976; Schreiber, *et al.*, 1978). The SIMED model accurately predicts the growth of alfalfa, as measured by above-ground dry matter accumulation. This modeling approach could be used to predict the performance of a CELSS, and provide a means of evaluating alternative paths for correcting off-course conditions. By monitoring state variables such as rate of photosynthesis, transpiration, and leaf area, actual conditions can be compared to the simulated values which are considered to be the true course.

Expert Diagnosis of Disorders

Latin, *et al.*, 1987 developed an expert system for diagnosing disorders of muskmelons and other cucurbits. By asking a series of questions, and evaluating each one through a hierarchy of if-then rules, the expert system was able to quickly reach a conclusion concerning the condition of the crop. If this approach, which requires the user to describe visual symptoms of the disorder, could be combined with the machine vision and image processing intelligent sensor developed by Hetzroni and Miles (1992), then an on-line approach to determining off-course conditions could be developed. An expert system combined with an intelligent sensor would be able to detect problems like nutrient deficiencies earlier than a human would observe the phenomena. By detecting and correcting problems early, while they are still manageable, corrections to get the CELSS back on-course can be small and less dramatic.

AUTOMATED WHEAT PRODUCTION IN A CELSS

Because the potential scenarios for a CELSS are so complex and varied, it is almost impossible to conceptualize automation scenarios without considering a specific example first. By studying how wheat may be grown in a CELSS, it is hoped that the general requirements for automation will become evident.

Planting

Planting is the process of transporting seed or propagule from the storage area, opening the container, and placing the seed in the desired location, at the proper depth, orientation and spacing.

Process Model

The prime requirement of planting is to ensure that the seed or propagule is in contact with the soil, or wicking apparatus, that provides adequate exchange of essential nutrients, primarily oxygen and water. During the germination process, the seed swells as it imbibes water, absorbs oxygen, and respires CO_2. Germination also requires darkness and a proper range of temperatures.

Current Practices

Hydroponics and nutrient film techniques have been devised for growing plants in a CELSS. At JSC nutrients are circulated through trays which have fiberglass wicks inserted in rows. Imbibed wheat seeds are sown in the rows between adjacent wicks. As the seeds germinate, roots extend down into the tray and form a mat that absorb the essential nutrients from the liquid solution. By imbibing the seed, more uniform stands can be established, but since soaking softens the seed coat, this technique requires much easier handling to avoid damaging the emerging tissues.

Minnesota basalts have been ground to particle size distributions which simulate lunar soils, and used as media for growing plants. In this process, the lunar simulant is spread evenly across trays through which water with added nutrients is pumped. Seeds are planted directly in the simulant.

Dreschel, *et al.*, 1988 have devised a method of circulating nutrient solutions in porous tubes under slight vacuum to prevent dripping. Seeds or seedlings are placed on the tube, and wrapped with black plastic sheath to shield the roots from light. A plastic tube cut along its length is placed over the plastic to hold it and the roots against the porous, nutrient supply tube. The roots wick the nutrients from the solution by capillary action across the porous tube.

In a demonstration of a commercial robot's capability, Boeing personnel at Kennedy Space Center (Parker and Eckhoff, 1989), equipped a robot with a suction tip end-effector, and programmed it to dip the tip into a canister of seed, then when a seed had plugged the suction hole, transport the seed over to a seed tray. While this scenario results in the simplest end-effector, it is very slow. A multiple-tip end effector would seed the flat much faster (del Castillo, 1987), but adds complexity and mass to the design.

Commercially, wheat is usually seeded with drills with disc openers about 7 inches apart. For seeding rates of 1 bushel per acre and seed counts of approximately 15,000 per

lb., the within drill spacing should be about 1 seed per drill-inch. A crop similar to wheat, rice is usually transplanted. Rice seedlings are started by germinating the seeds in a flat without cells. Workers or machines then pinch plugs out of the mat of roots and shoots and place each one individually into the soil. Often, rice is transplanted directly into puddled soil and flooded with water.

In the bedding plant industry, a considerable amount of automation has been developed (Chen *et al.*, 1992; Gautz and Wong, 1992; Honami, *et al.*, 1992; Kondo, *et al.*, 1992; Mohapatra, *et al.*, 1992; Morimoto, *et al.*, 1992; Nambu and Tanimura, 1992; Onoda, *et al.*, 1992; Roberts and Swanekamp, 1992; Sakaue, 1992; Shaw, 1993; Suggs, *et al*, 1992; Tanaka, 1992. Customary practices include sowing seeds into flats with many small cells filled with a light-weight soil mix, and transplanting the seedlings into a grow-flat where the plants continue to grow. By germinating seeds in small cells, better control of environmental conditions is possible and this leads to healthier plants and a higher percentage of germination. Because the seedling flat cells are 10% the size of the grow flat, transplanting frees up some greenhouse space for a few days. Since 100% germination is practically impossible to obtain, even with selected seed, the seed flats often have 20% misses or more. During the transplanting operation, these empty cells are skipped, and the grow flats 100% populated. By transplanting, the growth chamber, or greenhouse space is better utilized.

Seeding is often accomplished with a rotating, perforated drum on which a vacuum is pulled. The drum constitutes the bottom or side of a seed hopper, and by rotating upwards, seed are sucked against the tiny holes and carried up and out of the hopper. On the opposite side, the seed are released into tubes which route them to a row of cells in the flat. The drum rotation and flat advancement are coordinated so that one seed is placed in each cell. Usually the seed are then covered with a small amount of vermiculite or soil mix. Because the seed are dropped, they often roll to the corners of cells, which make them difficult to transplant mechanically.

Because each plant species has unique requirements for root and canopy environments, a myriad of seed flats and grow flats have been developed by commercial growers. The size, shape, depth, density and pattern of cells vary greatly. Some are in neat rows and columns, others are arranged to facilitate movement of air in the root zone and canopy. Because of this diversity, automation of a full range of tray configurations has not been possible until the introduction of robotic transplanters such as ROTRAN® 2000.

Proposed Automation

The tasks of transporting the seed from storage, opening the container, singulating the seed, and placing the seed in the growth media at the proper spacing, depth, and orientation can be automated by a number of suggestions.

Pencil Seeder

One possibility is to develop a storage container which also performs the singulation task, and which can be grasped by the end-effector and positioned by the robot to the proper locations for planting each seed directly in each grow-flat. The robot would supply the power (pneumatic, vacuum, or electric) required to operate the singulation

mechanisms. This device is envisioned as a large pencil, in which the barrel holds the seed, and the point of the pen is used to hold the individual seed and place them into the proper places in the grow-flat. The singulation might be accomplished by a ball-point pen type of mechanism. A push on the top of the pen would cause the device to acquire another seed from the barrel, and leave it on the tip ready for insertion into the lunar simulant, or placement on a wick. The pencil seeder has several attractive aspects:

- Planting the seed in the same container used for storage eliminates a materials handling task and simplifies the process of seeding;
- The pencil seeder can be quickly adapted to any type of robot;
- In case of robot failure, the pencil seeder becomes a tool used by an astronaut to mechanize planting and to reduce the labor requirements; and
- The pencil seeder has considerable commercial spin-off for small farmers, especially producers of specialty crops where seed quantities are small and seed prices often reach several thousand dollars per pound. The pencil seeder actually favors the small business because it handles the small quantities required in a CELSS.

Seed Tapes

Tapes in which seeds are placed at the proper intervals could be considered for automating the seeding process. The seed tape would be stored until needed, then the proper amount unrolled and placed on the growth chamber trays. This would require the robot to spool and cut the tape to the proper length for each row. End-effectors to accomplish this task would be simple and easy to design. The tape would be made of a wicking material that provides the proper flow of nutrients to the germinating seed. For missions where all the seed are supplied from Earth, the material could be bio-degradable. Eventually, the material needs to be reusable, or if biodegradable, made from plant materials generated as a by-product, such as rice, cotton, or linen fiber mats. The features of this technology are:

- The end-effector mechanisms to spool and cut the ribbon are simple and easy to design;
- Commercial technology already exists to place seed into seed-tapes prior to a mission; and
- The technology could be readily adapted for manual use if the robot fails;

Gels or Foams

Seed could be prepackaged with dehydrated gels or foaming agents in containers similar to tubes of caulk which can be squeezed, or pushed. By injecting water several hours prior to planting, the seed could be pre-germinated. After adding water, the tubes would need to be stirred, or rotated to insure uniform distribution of seeds in the gel. The robot would grasp the tube, and move it across the trays, depositing the correct amount of gel or foam. For some, high-density applications, the robot might lay down a continuous bead of gel, while in other cases, the seed would be placed in hills or squirts. The plant

population would be determined by the statistical density of seeds in the gel, and the volume of gel applied, and the positioning of the robot manipulator. Foams are mentioned with gels because the carrier material must provide not only liquid nutrients, but for rapid exchange of carbon dioxide and oxygen as well. This approach has several attractive features:

- Gels are a commercial method of planting high-value seeds and much of the technology is readily adaptable;
- The sensing and controls required for robotic application of glue and weather-stripping have already been worked out and can be quickly adapted for this use;
- In case of robotic failure, an astronaut could wield a "caulking-gun" filled with a tube of gel seeds;
- Seed are stored in the application container, saving a materials handling step;
- The knowledge gained by researching the nutrient transfer characteristics of gels and foams would have immediate impact in commercial gel products; and
- This approach to seeding could also be marketed to small, family producers of high-valued seed crops.

Seed-Flats and Transplanting

In order to optimize the use of the growth chamber, seeds could be planted into seed flats with small cells, then transplanted into pots or grow-flats after the seedlings have outgrown the seed-flats. For the germination and seedling growth stages, the seed flats would occupy about 10% of the space required by the fully-grown plants. For a few weeks at least, the seed-flat or transplanting approach would make growth chamber space available for other crops. The technologies required to place the seed into the seed-flats could be any of the above-mentioned ones, including the commercial, drum-style. In this case, the seed-flat would be brought to a stationary seeder, instead of taking the seed to the grow-flat located in the chamber.

Once the seedlings have filled out the space available in the seed-flats, they can be transplanted to the grow flats or pots. During the transplanting process, inferior seedlings can be removed, thus only the best plants would occupy space in the grow-flats. Grading could be based on superior growth rates, leaf areas, stem diameters, etc. The need to do this comes from the requirement to make optimum utilization of space in the growth chamber. Because of genetic variation even within a cultivated variety, individual plants may have several times the productivity of "average" plants. As a result, this technique could significantly increase yields.

The seed-flats should be constructed of porous materials which enable nutrients to flow freely to the seed. New designs will be required since commercially-available, plastic flats are practically impermeable to liquid and air flow, and rely on a drain hole in the bottom of each cell. The seed and grow flats proposed here would be made of fiber-glass, or porous plastic material that wick nutrients from beneath the flat, and provide adequate air exchange. The seed and grow flats would sit in nutrient delivery trays through which liquid media and oxygen are pumped. Control of the nutrient delivery system would

provide for ebb and flow to prevent saturation or drowning of roots. The soil media used in the seed flat should be finer than the grow flat, and can be made from decomposed plant materials, or lunar or Martian materials.

The seed-flat or transplanting method has a number of salient features:

- Seedlings grow better in smaller, confined cells provided by the seed flat;
- Less space is consumed during the juvenile stages of growth, thereby freeing chamber shelf space for other crops;
- The seeding process can occur in a single location, with the seed-flats transported to the growth chamber; (Designing machinery to automate seeding may be much easier.);
- Higher crop productivities are likely because of plant selection during transplanting;
- Transplanting can occur into pots, making it possible to place plants or clumps of plants individually; Plants requiring higher light intensities could be placed where such conditions occur. Plants requiring low light levels could be placed on the edges, or sides away from the light. This could lead to inter cropping, and substantial increases in food productivity without increasing the chamber size;
- These technologies could pave the way for improving performance, lowering costs and raising profits of commercial greenhouse producers of plants. The potential for commercial spin-offs in the plant cell culture industry is highly significant.

Culture

The efficient production of wheat in an environmental chamber requires control of the nutrient supply and exchange rates. The level of nutrients must be between minimum required for growth but less than toxic concentrations. Environmental variables, such as temperature, humidity, radiation, and nutrient pH must be such that the rate of basic physiological processes is satisfactory to sustain growth and development of plant tissues. Since the processes are complex and dynamic, the model must consist of a set of numerical equations describing the rates of material transfer as functions of physiological states and environmental conditions. As much as possible, the equations should be cause and effect models, to permit them to be used to predict responses for future conditions.

Numerous researchers have collected data on wheat growth in controlled environment chambers and developed models of the results. Work by J. T. Ritchie and S. Otter (1985, and 1987) has resulted in the CERES-Wheat model. Recently, Canpolat and Bolte (1993) converted CERES-Wheat into an object-oriented model that facilitates updates and changes, and makes it more usable as a management tool. Salisbury and Bugbee (1988a,b; 1989; 1991) showed that the potential production is far greater than record yields achieved in the field. Volk and Cullingford (1989) and Volk and Rummel (1987) developed BLSS, a model of wheat that tracks the flow of carbon, hydrogen, oxygen and nitrogen through the complete processes of a CELSS.

Current Practices

Wheat has been grown hydroponically in environmental chambers at Kennedy and Johnson Space Centers. In each case, seed are pre-soaked, then placed on wicks which supply nutrients from a liquid solution pumped into the trays beneath the wicks. The seeds germinate, and the roots extend into the tray, forming a mat which take-up nutrients and water. By the end of the growing season, the roots have filled the entire rooting zone. Wheat has also been grown on the porous tubes developed at KSC.

Wheat is grown in controlled environments only for research purposes due to the considerable cost difference over field practices. As a result, there is little mechanization of a scale suitable for a CELSS. However, practices developed for other crops may be adaptable.

Because lighting is a large energy cost and radiation not absorbed by plant leaves is wasted, relocating and spreading of plants is a common practice in greenhouses. Plants grown in pots are moved apart as the leaf area increases to absorb additional radiation. In some commercial facilities for lettuce, plants are grown in trays that move progressively apart as the plants mature. Previously proposed scenarios for CELSS would have a system of conveyor-tables which moves every few minutes to expose all the plants to the same amount of light and the same photo period. In a simple system, where the return conveyor is beneath the illuminated tables, the lights would stay on 24 hours per day, but each plant would receive only 12 hours of light, unless a supplemental bank of lamps were placed beneath the top conveyor to shine on the bottom tables. The motion of the plants would also dampen out spatial differences in light intensity. All such systems seek to have leaves intercept 100% of the light.

Another concern for illumination is the peak energy consumption. By adjusting the timing of the light period for each shelf, the peak can be greatly reduced. For an 8 hour photo period, the peak would be 1/3, and for a 12 hour period, the peak would be 1/2. For photo periods greater than 12 hours, the peak would be the same, but the area (power or wattage) would be reduced.

Proposed Automation

The one approach that appears to provide the versatility in environmental conditions required for the CELSS-candidate crops is ebb and flow tables with pots and/or flats. The volume of each pot must accommodate the root mass at plant maturity. The top surface of the pot must be as small as possible, to permit plants when small to be close enough to intercept the maximum amount of light. Then, as the plant leaf area enlarges, the pots can be spaced apart. The pot would rest on ribs or fingers formed in the bottom of the table, to facilitate drainage and air exchange. The nutrient solution would be pumped onto the table to a predetermined level that permitted the soil or wicks in the pots to saturate and wet the plant roots. Then the nutrient solution would be drained from the table, which would permit air to replace the liquid in the soil mix or wicks, much like natural, field conditions. Thus, the name ebb and flow. Although there are suggestions that hydroponics techniques are more efficient in carbon utilization than soil techniques, there is no reason that if provided similar rates of nutrient flow to the roots, that one system would grow a root to shoot ratio different from the other. The root mass (carbon) typically tied-up in a soil mixture can be extracted from an inert simulant or lunar soil by supersaturating it and developing a fluidized bed (quicksand), then lifting the plant by its

shoot. Once the plant was extracted from the pot, the soil mix could be dumped, washed, and reprocessed for the next crop. The ebb and flow technique should air-prune the roots (Huang and Ai, 1992; Chun and Takakura, 1992), and cause less root entanglement with the soil particles. But if not, then the soil can be recovered by oxidizing the root mass, which would also recover carbon and other nutrients. This would obviously be much simpler to automate than hydroponics which requires wicks to be cleaned and reprocessed between crops.

This approach would require that carriers be developed to transport the pots from the central processing station to each growth chamber table. The carriers would be held by a revolute-joint robot that moves along a rail in each chamber. At the airlocks between chambers, the robot would secure the carrier, and then affix itself to a special joint in the next chamber, which would permit it to be powered as it detached itself from one rail and attached to the rail in the next chamber. This capability would permit the robot to move along rails, yet cross obstacles such as airlocks and doorways. And, it would permit the robot power leads to be the length of one rail (probably one chamber long). The same procedure would be followed when going from one level in a chamber to another. The rail would be suspended from the "ceiling" in each chamber. This robot would have a number of end-effectors which permit it to grasp different objects and perform the required tasks. For example, when spacing the pots in the ebb and flow tables, the robot would traverse the rail until coming to the bay requiring the plants to be spaced. Then it would reach out and grasp each pot to be removed and place them in a carrier. When the required number of plants have been moved, the robot would move each remaining plant to the desired place on the table, according to predetermined patterns. The plants removed from this bay would be transported to other locations in the growth chamber and positioned correctly on tables by the robot.

Although the scenario described above refers to stationary tables, there are advantages to moving them along a conveyor, to pass beneath a bank of lights, and various stations where the required tasks are accomplished. Vertical lifts at each end of the sliding (conveyor) shelf would transfer the rectangular table from one level to another. The advantage of this system would be to minimize the number of lights required, but at the cost of additional hardware to move the tables of plants. Usually, systems which do not require motion are more reliable and require less maintenance. Since a robot carrier will be necessary to transport materials to and from the preparation and processing areas, letting it also provide the movement of plants will reduce the complexity, size and expense of the growth chamber. And, it is not likely that the launch weight of the extra automation mechanisms for the conveyor tables would be less than the weight of the lights required to illuminate all the growth chamber shelf space.

However, a single moving conveyor would bring all the plants past a workstation at the end, thereby alleviate the need for an aisle, and make the growth chamber more efficient. In a cylindrical chamber, the shelves could rotate on a conveyor which moves around the axis of the chamber. The workstation or aisle space would be in the center and the shelves rotate around it. The lighting would be around the inside wall of the chamber. The aisle floor would block out light from the plants beneath it, creating a dark or night period. The height of the floor would be such that the plants receive up to 16 hours of light per day, with 1 revolution of the shelves per day. The ideal width of the shelf must

be determined from light interception models. It's possible that two or more conveyors could rotate concentrically, with plant species requiring lower light intensities on the inside. In either case, the shelves would be populated with plants in pots.

The potted plant approach has a number of desirable features:

- Multiple robots could perform simultaneous tasks, or be held in reserve in case of failure of the primary unit;
- By using pots, the seeding, transplanting, harvesting, and recycling of pot materials would occur in a central work station, thereby simplifying the mechanization of materials handling;
- The energy efficiency would increase because pots can be easily spaced further apart when the plants grow large;
- Pots facilitate inter cropping which would accommodate variances in environmental conditions, particularly light intensity; and
- In case of failures, single pots can be replaced rather than entire trays, and the production quickly return to normal capacity.

Gathering (Harvesting)

As wheat matures in the field, it dries and turns a golden brown, due to physiological aging, and to stresses imposed by high temperatures and late Spring droughts. In some situations, wheat and rice are cut and left to dry for a few days before combines (combined harvesters and threshers) are used. This decreases the energy required to thresh and separate the grain. In the growth chambers, drying can be accelerated by withholding the nutrient solution, decreasing ambient humidity, and elevating air temperature.

The optimum time of harvest is a tradeoff between grain moisture content and the need to plant the next crop. The grain must be below 14% to minimize the risk of molds and to store well for up to 12 months (Brooker, *et al.*, 1992), but if allowed to dry in the chamber to 14% or less, the kernels shatter easily, leading to serious pre harvest and gathering losses, and possibly increasing damage to the grain during threshing. The optimum grain moisture content to minimize gathering losses is about 18%. Obviously, the temperature and humidity conditions must be such to not only facilitate the dry down process, but to accelerate it. Each day the crop stands in the growth chamber beyond physiological maturity is a day which could have been used to grow the next crop.

The translocation rate of carbohydrates to the seed, the physiological maturation of abscission layers , and the process of drying in the spikes is not a well documented aspect of wheat growth. Data are required to develop adequate mathematical expressions of the relationships, in order to simulate the process and determine optimum harvest strategies.

Current Practices

The time required to harvest 11.4 m^2 of wheat grown in the plant growth chambers at Johnson Space Center, place the above ground and root portions in bags, weigh the bags, and clean the trays was observed to be approximately 1.5 hours for 20

people. That's approximately 2.6 person-hours per m^2 of effort to harvest and initiate biomass recovery and processing of a crop which took 80 days to grow. That averages about 2 minutes per day. But with startup and cleanup, a crew member is likely to spend much longer if harvest occurs each day. A much better scenario would be to harvest more than one day's supply at a time. The exact amount will be a tradeoff between the demand for space in the growth chamber, labor requirements for harvesting, and the costs of storage of the processed grain.

The numerical values of such statistics have little long-term merit because no accurate measurement was made of the time and motion of each worker, or was each worker supervised to be sure they were constantly performing their task. The processes which are best suited for mechanization or automation may be considerably different, and require much different times. Comparisons should be made only after mechanization and automation are integrated into the processes.

At Kennedy Space Center, McDonnell-Douglas personnel undertook the task of developing the sensor and robot control programs for a robot to reach into a canopy of mature wheat and remove each head from the stalk. Such an approach would minimize the amount of biomass which must be threshed, but it proved to be a most difficult and time-consuming task.

Commercially, farmers harvest grain crops with combines (machines that combine the processes of gathering, threshing and separating) with field capacities of 5 or more acres per hour. At this rate, the 11.4 m^2 chamber area would be harvested in 2 seconds! The crop is cut with a sickle-bar and guided by a finger or bat type reel onto a cross-auger which carries it across the grain platform to the feeder-conveyor which takes the crop to the threshing and separation mechanisms. The grain platform headers are often 20 feet or more in width. By adjusting the height of the platform, the operator can cut the crop just below the heads, thereby reducing the volume of biomass which must be threshed, which reduces the energy requirements for threshing and the amount of trash in the grain.

Japanese rice combines gather each row of stalks and by maintaining the plant orientation are able to insert just the heads for threshing. This procedure minimizes the energy required for threshing, it keeps the rice straw intact, whole and useful for mats, thatch, etc., and it permits rice to be harvested when the straw is relatively wet and green.

These large, American style machines have been designed for 1 G conditions and operate in only a very narrow range of crop properties, particularly moisture content. The scale of these machines is certainly inappropriate, and probably the materials handling mechanisms (cutter-bars, augers, conveyors, etc.) are as well. The Japanese rice harvesters maintain a positive control of the plants after cutting, which may make this technology adaptable for lunar or Martian gravity.

Proposed Automation

Gathering the wheat crop would be greatly facilitated by growing the plants in pots which can be quickly removed from the growth chamber to initiate the dry-down process. The pots and soil/wick mixture can be immediately processed and recycled for the next crop. The wheat crop (roots and all) can be held in ventilated racks where waste heat and low humidities are used to dry the grain to the desired level. The robot arm would perform the task of moving the pots into carriers and transporting the carrier to the pot

and soil recovery area, and after the plants are extracted, moving them to the dry-down area. Because of the versatility of the robot, its grippers, and its sensors, very little if any new hardware will be required for gathering. The programs will be different from planting.

Threshing and Separation

Once the crop has been gathered from the field and dried to the best moisture content for threshing, the grain kernels must be detached from the head and separated from the chaff or other plant parts. Detaching requires mechanical, rubbing motion with a minimum of impact which could cause damage to the grain. Centuries-old practices for separation (winnowing) utilizing differences in particle densities and their aerodynamic drag coefficients are also followed by modern machinery.

A model of the soybean material flow through a combine was developed by Miles and Tsai, 1987. Materials were divided into categories: heads or pods (with kernels attached), free grain (kernels detached or free), and MOG (material other than grain). The equations describing the rate of material flow through the combine and from one category to another were based on bio-physical properties of the crop and the machine design parameters such as conveyor speed, and cross-sectional area. Reed Turner developed the HarvesTrainer® personal computer program which models the harvesting, threshing, and separation of corn, wheat and barley for several models of John Deere combines. Mathematical models show the complex interaction between doing an adequate but not excessive job of threshing, and the size of sieve opening and counter-flow air volume (Mailander and Krutz, 1984; Mahoney and Srivastava, 1986; Kim and Gregory, 1989; Bjork, 1991;Nath, et al., 1982; Trollope, 1982). Excess threshing not only detaches the seed, it grinds the heads and stems into fine particles that are difficult to separate from the grain. Sieve openings may be adjusted to screen out the larger chaff particles. The fan speed must be adjusted to provide a sufficient counter-flow of air through the louvers to exceed the terminal velocity of the chaff, but not the grain. Problems occur when the thresher breaks the head and stems into a particle size distribution whose aerodynamic drag coefficients and densities overlap those of the grain. By properly adjusting the machinery, and by threshing when the proper crop bio-properties exist, clean, damage-free grain can be obtained.

Current Practices

Researchers typically use a plot-thresher which detaches and separates the grain. In such a device, wheat plants are fed by hand in small quantities. A rotating drum with spikes rubs the head against a stationary set of spikes, and the detached kernels fall down into a catch pan. A small fan provides a cross-flow of air that blows the chaff into a bag.

Modern combines not only gather the crop, but also thresh and separate the grain as well. Sensors detect grain loss and measure the bio-properties of the crop, and microcomputers adjust the machine to optimize performance. Properly adjusted, modern combines harvest 95% or better of small grain crops. Over 80% of the losses occur at the header during the gathering process.

Conventional combines use rotating cylinders with rasp-bars, which rub the grain against open grates called concaves. Occasionally, in tough conditions, special rasp-bars

and concaves with intermeshed spikes are used. The clearance between the rasp-bars and concaves is adjustable for different crops and quantity of biomass. The openings in the concaves can be changed by adding or withdrawing curved wires or rods. The flow of material is radial, or tangential to the axis of rotation. As the heads are threshed, the kernels and chaff fall through the grates in the concaves to the cleaning mechanisms. Secondary separation occurs in the straw-walkers.

Axial flow combines use similar threshing mechanisms, except the flow of material is along the axis of the cylinder (termed a "rotor") which is mounted parallel to the axis of the combine. This mounting permits much longer and larger diameter cylinders (rotors) to be used. Since the threshing occurs over a much greater area, the process is usually more gentle than the conventional threshing mechanisms, and less grain damage occurs. Manufacturers claim that grain separation is enhanced by centrifugal force of the rotating cylinder.

Although the threshing and initial separation mechanisms of conventional and axial-flow or rotor machines are significantly different, the cleaning mechanisms still rely on the proven technologies of oscillating sieves with a counter-flow of air. Louvers in the sieves are adjusted to permit the grain to freely pass, but prevent larger biomass pieces to bounce across the top due to the oscillatory motion. Chaff particles the size of grain are prevented from falling through the sieves by counter-flow of air. The velocity of air through the sieve openings must exceed the terminal, or settling velocity of the chaff, but not that of the grain kernels. Air velocity is adjusted by modifying fan speed. Unfortunately, the actual velocity through the openings is affected by the density and uniformity of the mat of chaff on the sieves. The denser, heavier and thicker the mat, the greater the differential pressure, and the less the velocity. In case of non-uniformity, the thin areas may blow completely off, which permits most if not all of the air to flow through the hole. Excessive trash in the grain results from this situation.

Proposed Automation

The size of modern combines is obviously much too large for a CELSS, but the efficiency and reliability of the proven mechanisms makes them very attractive for consideration. As the plot threshers have proven, the mechanisms for threshing and separation can be scaled to an appropriate size. No doubt, a robot could grasp each bundle of wheat and place the tips containing the heads into the thresher. However, it is not clear that the separation and cleaning mechanisms (sieves and fans) will work properly in less than 1 G conditions.

A suggested alternative is to accelerate the threshed grain and chaff mixture and utilize the difference in particle momentum to separate the kernels from the chaff. The mixture could be forced pneumatically out a tube onto a slightly cupped, spinning disk with short blades (similar to the spreader mechanism used on a bulk, dry fertilizer truck). Coming off the spinning disk, the kernels would have the greater momentum, since they would have the greater mass. An opposed flow of air would halt the horizontal movement of chaff quickly because of their larger aerodynamic drag coefficients. The grain kernels would move further away from the spinning disk because of their greater momentum and less drag. Additional enhancements, if necessary would be to create a velocity gradient in the opposed flow of air. This would permit particles to settle out of the horizontal stream

at distances proportional to their coefficients of aerodynamic drag. Additional separation is possible by adding a spiral separator, which permits the heavier and more round particles to roll faster, and be carried out by centrifugal force. Vibratory separation should also be considered.

The clean grain should be placed in a container which the robot can carry to a temporary queue, place it the dryer, and/or place in storage. While the chaff residues from the separation process can be returned to the crop biomass, the particle sizes are now much smaller, so a container must be used. This material can be fed into the machinery for the recovery process.

Recovery and Recycling of Inedible Biomass

After the grain kernels have been detached, the stems and roots must be processed to recover the nutrients. Among the processes proposed are leaching, bio-digestion, and oxidation, all of which are more efficient if the plants are shredded or macerated. Leaching nutrients with acids and bio-digestion with enzymes, micro-organisms and/or animals may be better with the wet, green plant tissues. Oxidation would undoubtedly require not only fine particles, but dry materials as well.

In agriculture, devices used for changing the particle size of plant materials include forage choppers, hammer mills, and grinders. Forage choppers leave particles approximately 1 to 2 inches long, and work well with wet, green materials. Hammer mills and grinders work well with dry materials, and are capable of particle sizes of .1 inch or smaller. Hay conditioners which scuffs the outer layer of the stems and/or crimps the stems, are able to speed the drying process for the biomass.

The materials handling requirements to feed the plant residues to such machinery are rather simple. If the crop is still bunched, a robot could grasp the clump and force it into the machine. If not, the crop can be placed onto a conveyor which accelerates it to thin it out to the desired thickness and feeds it into the machine. The output of any of these machines must be delivered to the apparatus which effects the nutrient recovery. This could be by mechanical conveying, pneumatic conveying, or containers carried by robots.

Preprocessing

The most likely preprocessing of wheat is drying the moisture content below what is best for harvesting (18%) to what is best for storage (14% or less). For storage less than a year 14% is considered acceptable for up to 2 years it is 13%, and for 5-years the wheat should be dried to 11-12%. Usually this is accomplished by heating air to reduce the relative humidity and blowing it through the grain (batch-dryer). Alternative methods include flowing grain across heated metal plates (continuous-flow dryer). Occasionally, for small batches of grain, microwave ovens are employed. Too rapid drying creates stress cracks in the seed whereas drying too slowly permits microorganisms to grow and damage the quality of the grain. If the drying time is too long, more moisture than necessary is removed, and in extreme cases, the seed become "cooked". The proper moisture content, drying rate, and drying time also depend on what the grain is used to make (Bruce, 1992). It is especially important to note that whenever any of the grain is

kept as seed for subsequent crops, that the damage during drying be kept low to maintain viability.

The basic principle is to keep the equilibrium relative humidity for the grain below 68%. Safe storage conditions can be determined from an equilibrium moisture content curve as given in ASAE Data D245.4 (ASAE, 1993) which are different for hard (Durum, high-protein, used for bread) and soft (pastry) varieties. Numerous models have been developed to predict the drying time and quality of grain, especially corn (Chung and Verma, 1991; Bunn and Wishert, 1991; Bruce, 1992; Giner, *et al.*, 1991; Parti, 1990; Sanderson, *et al.*, 1989; Sokhansanj and Bruce, 1987; Abawi, 1993).

The amount of grain harvested at one time from a CELSS plot is relatively small and can easily be placed in drying ovens by a robot. The ovens used to dry the biomass may also be used for drying the grain, provided the temperature is as prescribed by the ASAE procedures. The grain should be uniformly scattered across a pan with a screen for the bottom, and placed on a shelf in the oven. The oven should have air flowing through the screen vertically to remove the moisture. The sensors in commercially-available, hand-held grain moisture meters can be adapted to provide on-line, continuous measurement of moisture without opening the oven. Another sensing technique would be to monitor the air above the grain pan to detect a rapid decline in relative humidity, which would signal that the grain has reached the equilibrium moisture content.

Mechanization for this process requires that the container of grain from the threshing and separation process be poured uniformly on the drying pan, which is then placed on a shelf in the oven. After the grain is dried, the dying pan is removed from the oven, the grain is poured into a container which is sealed and placed in storage. The same robot used to manipulate materials in the growth chamber can also be used for these tasks, provided the base-rail extends into the processing chamber.

Storage

The design considerations for storage include the volume required for the raw, edible products; and the environmental conditions (temperature, humidity, and oxygen concentration) necessary to sustain the quality for the intended shelf-life. The containers used to carry, dry and store the grain must be efficient in utilization of storage volume. The storage volume required depends on the crop yield, the harvest area, and the food reserve factor. For example, if 11.4 m^2 of wheat yielding 100 bu/A were harvested, then approximately 0.35 cubic feet of storage would be required. If a 3x reserve of wheat is desired, then 3 x 0.35 = 1.05 Ft^3 would be needed. Please note that two of the storage bins would normally be full, but the third would be somewhat less than full. Thus, the true amount of reserves is one less than the number of storage sites.

In storage, the grain moisture content will reach an equilibrium with the surrounding conditions, depending on temperature and humidity. Seeds very slowly respire, which converts stored carbohydrate and oxygen into carbon dioxide and water. Respiration proceeds with a Q_{10} of 2; that is, it doubles for each 10° C rise in seed temperature. Low temperatures also inhibit the growth of microorganisms such as molds and bacteria that lower the quality and in some cases make the grain unfit for human consumption. Thus the ideal conditions for grain storage are low temperatures (even below the freezing point of water), and low levels of atmospheric moisture and oxygen.

The automation required for storage includes automatic insertion and retrieval of containers of food. This could be on shelves, 1 unit deep, on which the robot could place items. A more sophisticated approach would be to develop a miniature warehouse storage and retrieval system with inventory control. Arguments for the robotic approach are similar to those used for the robotic handling of plants in the growth chamber. With machine vision, the robot could "see" each vacant slot and insert the food container into the next available space.

Food Processing

Just as an army travels on its stomach, the amount of work accomplished by an astronaut crew will depend on having a wide variety, and ample quantities of tasty, nutritious foods. When selecting crop species and food processing techniques to provide dietary needs, taste preferences and nutritional demands for humans working in space environments must be considered. Its probable that the loss of calcium and muscle during micro gravity conditions will require diets considerably different from earth-base recommendations. After arriving on orbit, astronauts seem to choose spicier foods to consume. To ensure needs are met, astronauts must be included in decisions about crops and processing.

The automated food processing devices must retrieve the raw food products from storage, measure and pour the desired amount into the hopper for the food processor, whether it be a milling machine, pasta machine, or toaster. Bread, pastas, and cereals require different processing: grinding, milling, cracking, toasting, and extrusion. The machinery for these processes are different, so to provide a variety of foods from wheat or other grains, several machines will be required. However, the concept of a grain container that can be poured into a hopper by the robot, permits the materials handling to be automated without a lot of complication. The difficult task is to clean the machinery after use. This will require special end-effectors on the robot which can disassemble the key components of the machinery and place them in a washer. Later the robot can retrieve the components and reassemble them.

Food Preparation

After the grain has been processed, it must be prepared into the food to be eaten by the astronauts. Flour that has been ground must be sifted, mixed with water and other ingredients and baked into bread, biscuits and rolls. Like other aspects of a CELSS, questions arise: How much of this should be automated? What tasks are better left to the crew? Should bread be baked in small amounts each day, or in larger batches each week? The answers probably depend on which tasks the astronauts enjoy doing, and which become drudgery to perform. As the frequency of occurrence increases and the drudgery becomes higher, the greater the need for automation.

Again the processes can be automated by robotic handling of containers of the food products.

CONCLUSIONS

The myriad of tasks required to grow plants, process edible portions into food, and recycle inedible biomass into nutrients required by subsequent crops, could require large amounts of astronaut labor unless automated (Schwartzkopf and Brown, 1991; Schwartzkopf, 1991; Schwartzkopf, 1993). Conceivably, the amount of labor required for performing all the materials handling for planting, tending, gathering, separation and threshing, processing, recycling wastes and cleaning equipment could exceed the life-support capacity of a CELSS. Thus it is essential that automation become an integral component of CELSS research. Because of their versatility, robots offer an overall less complex and less cumbersome solution to mechanizing the materials handling tasks than hard, fixed automation. Individual processes such as threshing, chopping and grinding are best accomplished by special purpose mechanisms. Researchers around the world are developing the sensors, end-effectors, manipulators and robot control programs necessary to automate materials handling tasks for typical, earth-based plant production. Some of the robotic transplanting, culturing, and harvesting efforts are applicable to a CELSS, but in most cases additional technologies must be engineered.

Techniques for growing crops in controlled environments are being developed by CELSS researchers. Data on oxygen and water recovery, carbon dioxide scrubbing, nutrient and energy input requirements, and food production are being gathered for a number of candidate crops. Techniques for germination and growth with hydroponics are being discovered. Failure to simultaneously develop mechanisms which not only work well with the bio-physiological processes but also automate the procedures and reduce the manual labor requirements is a serious oversight. Engineers charged with automating current, proven production and recycling practices will face enormous challenges unless cooperation begins with the biological scientists immediately. The bio-physical science research and the engineering for automation should proceed cooperatively. Working together, the scientists and engineers will be able to develop hybrid techniques which satisfy the biological requirements for life support and the operational constraints of space, launch weight and labor. This is the only way to insure the sucessful development of a CELSS.

Because the options for automation are so numerous, a general purpose solution suitable for all crops is difficult to conceive. A systems engineering study based on animated simulation of specific CELSS scenarios should be undertaken to evaluate and compare alternative designs. Which tasks should be automated by fixed engineering, and which ones should be automated by programming a robot to provide the necessary actions? Which tasks should be performed by the crew? How much time does it take for the robot to perform a task versus an astronaut performing it? How much power is required by the robot? How much space is occupied by the automated machinery? What does it weigh? If the volume and weight of a CELSS plus its automated machinery and processing area were used to store food, and if water and oxygen were recycled by physical-chemical means, how many people-days of life could be supported? The answers to such questions can only be obtained by following the systems study with a laboratory study to validate the proposed automation, and to collect statistics on human/machine interactive performance. In addition to answering a number of basic biological questions,

the proposed Human Rated Test Facility (HRTF) at JSC should be used to answer many questions concerning mechanization and labor requirements for a CELSS.

A basic concept for automating materials handling required to grow wheat in a CELSS is proposed to consist of removable benches on shelves, filled with containerized plants (pots), which are transported by a robot to and from the processing area. The robot rides on a rail mounted overhead, and has numerous end-effectors (grippers) which enable it to perform many different tasks at any location in the growth and processing chambers. Liquid nutrients are recirculated to the benches by ebb and flow techniques used in the commercial plant production industry. Pots enable the robot to space plants dynamically as they grow to utilize the maximum amount of light possible, to cull plants not performing to minimum expectations, and to replace the culled plant with a vigorously-growing new pot.

In summary, the conclusions are:

1. Materials handling in a CELSS must be automated,
2. Robots are superior to fixed-automation,
3. CELSS tasks require unique sensors, end-effectors and manipulators for robots,
4. The bio-physical science research and the engineering for automation should proceed cooperatively,
5. An animated simulation approach is necessary to evaluate the myriad of alternatives for automation.

Immediate funding of scientific and engineering efforts to automate materials handling will produce short-term benefits and help ensure the long-term success of Controlled, Ecological Life Support Systems when needed for exploration of the Moon, Mars, or the rest of our Universe.

REFERENCES

Abawi, G. Y. 1993. A Simulation Model of Wheat Harvesting and Drying in Northern Australia. Journal of Agricultural Engineering Research 54(2):141-158.

American Society of Agricultural Engineers. 1993. Standards. The ASAE, St. Joseph, MI. 49085-9659.

Benady; Meny, Yael Edan, Amots Hetzroni, and Gaines E. Miles. 1991. Design of a Field Crops Robotic Machine. Paper No. 91-7028. The ASAE, St. Joseph, MI 49085-9659. 7p.

Benady, M. and G. E. Miles. 1992. Locating Melons for Robotic Harvesting Using Structured Light. Paper No. 92-7021. The American Society of Agricultural Engineers. St. Joseph, MI. 49085-9659 14p.

Benady, M., J. E. Simon, D. J. Charles, and G. E. Miles. 1992. Determining Melon Ripeness by Analyzing Headspace Gas Emissions. Paper No. 92-6055. The American Society of Agricultural Engineers. St. Joseph, MI. 49085-9659. 13p.

Bjork, A. 1991. A Three-Dimensional Arithmetic Model to Calculate Grain Separation and Losses for a Rotary Combine. Canadian Agricultural Engineering 33(2):245-253.

Brooker, Donald, Fred Bakker-Arkema and Carl Hall. 1992. Drying and Storage of Grains and Oilseeds. AVI/Van Nostrand Reinhold. Florence, KY.

Bruce, D. M. 1992. A Model of the Effect of Heated-Air Drying on the Bread Baking Quality of Wheat. Journal of Agricultural Engineering Research. 52(1):53-76.

Bugbee, B. G. and F. B. Salisbury. 1988. Exploring the limits of crop productivity: I. Photosynthetic efficiency of wheat in high irradiance environments. Plant Physiol. 88:869-878.

Bunn, J. M. and C. A. K. Wishert. 1991. Adding a Wetting Routine to Thompson's Model. Transactions of the ASAE 34(4):1892-1899.

Canpolat, Necati, and John P. Bolte. 1993. Object-Oriented Implementation of the CERES-Wheat Model. Paper No. 93-4052. The American Society of Agricultural Engineers, St. Joseph, MI. 49085-9659. 18p.

del Castillo, E. Lopez. 1987. Robotic Automation in the Controlled Ecological Life Support System (CELSS). IN: Research and Technology Annual Report. John F. Kennedy Space Center. NASA Tech. Memo. 100303. pp 36-37.

Chen, S., Y. N. Chu, and C. F. Chang. 1992. Development of a tray stacker and destacker for vegetable seedling production. IN: Transplant Production Systems. Acta Horticulturae. 319: 517-522.

Chun, Changhoo and Tadashi Takakura. 1992. Dynamic Management of Root Exposure to Air Above the Nutrient Fluid Surface for Production of Tomato (*Lycopersicon esculentum* Mill.) Transplants. IN: Transplant Production Systems. Acta Horticulturae. 319: 483-488.

Despain, Ronald R. 1988. Plant Growth Machine Vision Development. Feasibility Study. McDonnell Douglas Astronautics Co., Design Engineering. John F. Kennedy Space Center, FL

Dreschel, T. W., J. C. Sager, and R. M. Wheeler. 1988. Status of Porous Tube Plant Growth Unit Research: Development of a Plant Nutrient Delivery System for Space. Paper No. 88-4524. The ASAE. St. Joseph, MI. 49085-9659. 15p.

Edan, Yael and Gaines E. Miles. 1991. Design of an agricultural robot for harvesting melons. Paper No 91-7029. The ASAE. St. Joseph, MI 49085-9659. 23p.

Edan, Yael, Bernard A. Engel, Gaines E. Miles, and U. M. Peiper. 1991. Intelligent Control of a Robotic Melon Harvester. *Revue d' automatique et de productique appliques* 3(3-4):37-53.

Edan, Y. and G. E. Miles. 1993. Design of an Agricultural Robot for Harvesting Melons. Trans. of ASAE 36(2):593-603.

Gautz, L. D. and C. K. Wong. 1992. Automation of the Handling and Manipulation of Vessels During Micro propagation. IN: Transplant Production Systems. Acta Horticulturae. 319: 573-578.

Gill, Paul S. 19__. FARMS: The Flexible Agricultural Robotics Manipulator. pp344-349.

Giner, S. A., C. E. Lupano, and M. C. Anon. 1991. A Model for Estimating Loss of Wheat Seed Viability During Hot-Air Drying. Cereal Chemistry. 68(1):77-80.

Guyer, D. E., G. E. Miles, M. M. Schreiber, O. R. Mitchell and V. C. Vanderbilt. 1986. Computer Vision and Image Processing for Plant Identification. Trans ASAE 29(6):1500-1507.

Hetzroni, A., G. E. Miles, P. A. Hammer, and R. X. Latin. 1992. Machine Vision Monitoring of Plant Health. Proc. World Space Congress, 29th Plenary Meeting of the Committee on Space Research (COSPAR). Washington, DC.

Hetzroni, A., G. E. Miles. 1992. Machine Vision Monitoring of Plant Health. Paper No. 92-3574. The American Society of Agricultural Engineers, St. Joseph, MI. 49085-9659. 12p.

Holt, D. A., G. E. Miles, R. J. Bula, M. M. Schreiber, and R. M. Peart. 1976. SIMED, A Crop Production Model, as a Tool for Teaching Crop Physiology. Journal of Agronomic Education (5):53-56.

Honami, N., T. Taira, H. Murase, Y. Nishiura, and Y. Yasukuri. 1992. Robotization in the Production of Grafted Seedlings. IN: Transplant Production Systems. Acta Horticulturae. 319: 579-584.

Hoshi, T., T. Abe and K. Nuki. 1992. Development of an Expert System Using Image Database for Diagnosing Plant Protection.

Huang, B.K. and F. Ai. 1992. Air-Pruned Transplant Production System for Fully Automated Transplanting. IN: Transplant Production Systems. Acta Horticulturae. 319: 523-528.

Iwao, K. and T. Shibata. 1992. Development of an Automatic Plant Growth Monitoring System Based on a Personal Computer. IN: Transplant Production Systems. Acta Horticulturae. 319: 613-618.

Kim, S. H. and J. M. Gregory. 1989. Power Requirement Model for Combine Cylinders. Paper No 89-1592. The ASAE. St. Joseph, MI. 49085-9659. 11p.

Kobayashi, T., M. Inagaki, S. Hata and M. Takai. 1992. Crop-Row Detecting System by Neural Network. IN: Transplant Production Systems. Acta Horticulturae. 319: 647-652.

Kondo, N., Y. Shibano, K. Mohri, T. Fujiura, and M. Monta. 1992. Request to Cultivation Method from Tomato Harvesting Robot. IN: Transplant Production Systems. Acta Horticulturae. 319: 567-572.

Kutz, L. J., G. E. Miles, P. A. Hammer and G. W. Krutz. 1987. Robotic Transplanting of Bedding Plants. Trans. ASAE 30(3):586-590.

Latin, R. X., G. E. Miles, and J. C. Rettinger. 1987. Expert Systems in Plant Pathology. Plant Disease 71(10):856-872.

Ling, P. P., and V. N. Ruzhitsky. 1992. Transplant Uniformity Inspection Using Machine Vision. IN: Transplant Production Systems. Acta Horticulturae. 319: 607-612.

Mahoney, W. T. and A. K. Srivastava. 1986. Property Based Combine Simulation Models. Paper No. 86-1578. The ASAE, St. Joseph, MI. 49085-9659. 25p.

Mailander, M. and G. Krutz. 1984. Development of a Dynamic Model of a Combine Harvester in Corn. Paper No. 84-1588. The ASAE, St. Joseph, MI. 49085-9659. 24p.

Miles, Gaines E. 1990. Robotics and Intelligent Machines. Agricultual Information Technology. Office of Technology Assessment. U.S. Congress, Washington, DC. 31p.

Miles, Gaines E. 1989. Plant Features Measurements for Robotics. IN: (Hosler and Armstrong, Eds) 1989 Research Reports. NASA/ASEE Summer Faculty Fellowship Program. John F. Kennedy Space Center and University of Central Florida. NASA Grant NGT-60002 Supplement 2. Contractor Report No. CR-166837. pp203-225.

Miles, Gaines E. 1991. Multispectral Image Processing for Plants. IN: (Hosler, Beymer and Armstrong, Eds). 1991 Research Reports. NASA/ASEE Summer Faculty

Fellowship Program. John F. Kennedy Space Center and University of Central Florida. NASA Grant NGT-60002 Supplement 6. Contractor Report No. CR-189744. pp 236-260.

Miles, G. E. and L. J. Kutz. 1991. Robots in Transplanting. IN: Bajaj (Ed). Biotechnology in Agriculture and Forestry. Vol 17 High-Tech and Micropropagation I. Springer-Verlag. pp 452-468.

Miles, G. E. and You-Jen Tsai. 1987. Combine Systems Engineering by Simulation. Trans. ASAE 30(5):1277-1281.

Mohapatra, Subhas C., and Charles W. Suggs. 1992. Optimization of Containerized Transplant Production: The Hybrid-Float System. IN: Transplant Production Systems. Acta Horticulturae. 319: 529-534.

Morimoto, T., S. Takahashi, K. Matsubara, and Y. Matsuoka. 1992. Development of an Automatic Propagation System of Seedling. IN: Transplant Production Systems. Acta Horticulturae. 319: 547-550.

Myjak, Michael, Mike Sklar, Roy Tharpe, Mark Thomas, Dan Wegerif. 1990. Automated Plant Growth Development Project. Advanced Technologies Branch, McDonnell Douglas Space Systems Company, 1989 Annual Report submitted to Special Projects Branch, Design Engineering, John F. Kennedy Space Center.

Nambu,T. and M. Tanimura. 1992. Development of Automatic Transplanter Using Chain Pot for Vegetable Crops. IN: Transplant Production Systems. Acta Horticulturae. 319: 541-546.

Narimatsu, J., T. Fujishiro, T. Kawata and K. Tsuchiya. 1992. A Hydroponics System for Raising Spinach Seedlings. IN: Transplant Production Systems. Acta Horticulturae. 319: 493-498.

Nath, S., W. H. Johnson and G. A. Milliken. 1982. Combine Loss Model and Optimization of the Machine System. Transactions of the ASAE 25(2):308-312.

Onoda, A., K. Kobayashi and M. Suzuki. 1992. The Study of the Grafting Robot. IN: Transplant Production Systems. Acta Horticulturae. 319: 535-540.

Otter, S., D.C. Godwin, and J.T. Ritchie. 1987. Testing and Validating the CERES-Wheat Model in Diverse Environments. AgRISTARS Publ. No. YM-I 5-W407. NTIS, Springfield, VA.

Parker, F. C. and J. L. Eckhoff. 1989. Design, Development and Testing of Plant Growth Seed Planting Prototype for Robotic Applications. Boeing Aerospace Operations. Robotic Applications Development Laboratory. John F. Kennedy Space Center, FL

Parti, M. 1990. A Theoretical Model for Thin-Layer Grain Drying. Drying Technology 8(1):101-122.

Ritchie, J. T., and S. Otter. 1985. Description and Performance of CERES-Wheat: A User-Oriented Wheat Yield Model. IN: ARS Wheat Yield Project. ARS-38. NTIS, Springfield, VA. pp 159-175.

Roberts, William J. and William Swanekamp. 1992. Production of Plug Transplants at Kube Pak Corp. IN: Transplant Production Systems. Acta Horticulturae. 319: 473-476.

Sakaue, Osamu. 1992. Development of Automated Seedling Production and Transplanting System Using Robotics. IN: Transplant Production Systems. Acta Horticulturae. 319: 557-562.

Salisbury, F. B. and B. G. Bugbee. 1988. Plant Productivity in Controlled Environments. HortScience 23(2):293-299.

Salisbury, Frank B. 1991. Lunar Farming: Achieving Maximum Yield for the Exploration of Space. HortScience 26(7): 827-833.

Sanderson, D. B., W. E. Muir, R. N. Sinha, D. Tuma and C. I. Kitson. 1989. Evaluation of a Model of Drying and Deterioration of Stored Wheat at Near-Ambient Conditions. Journal of Agricultural Engineering Research. 42(3):219-233.

Sase, S., M. Nara, T. Okuya and K. Sueyoshi. 1992. Determining Seedling Characteristics Using Computer Vision and its Application to an Expert System for Grading Seedlings. IN: Transplant Production Systems. Acta Horticulturae. 319: 683-688.

Schreiber, M. M., G. E. Miles, D. A. Holt, and R. J. Bula. 1978. Sensitivity Analysis of SIMED. Agronomy Journal 70:105-108.

Schwartzkopf, Steven H. 1991. Prioritizing Automation and Robotics Applications in Life Support System Design. Paper No. 911398. SAE, 400 Commonwealth Drive, Warrendale, PA. 15096-0001. 6p.

Schwartzkopf, Steven H. 1993. Design of a Controlled Ecological Life Support System. BioScience 42(7):526-535.

Schwartzkopf, Steven H. and Mariann F. Brown. 1991. Evolutionary Development of a Lunar CELSS. Paper No. 911422. SAE, 400 Commonwealth Drive, Warrendale, PA. 15096-0001. 9p.

Shaw, 1993. Plant Supply Systems for Automatic Transplanters. Paper No. 93-1021. The ASAE, St. Joseph, MI 49085-9659. 9p.

Shibata, T., K. Iwao and T. Takano. 1992. Growth Prediction of Lettuce Plants by Image Processing. IN: Transplant Production Systems. Acta Horticulturae. 319: 689-694.

Shimizu, Hiroshi and Minoru Yamaki. 1992. Non-Contact Growth Analysis Using Computer Vision System. IN: Transplant Production Systems. Acta Horticulturae. 319: 641-646.

Simonton, W. 1991. Automation in the Greenhouse: Challenges, Opportunities and a Robotics Case Study. HortTechnology 2(2): 231-235.

Simonton, W. 1991. Robotic End Effector for Handling Greenhouse Plant Material. Trans. ASAE 34(6):2074-2080.

Simonton, W. 1990. Automatic Geranium Stock Processing in a Robotic Workcell. Trans. ASAE. 33(6):2074-2080.

Simonton, W. and J. Pease. 1990. Automatic Plant Feature Identification on Geranium Cuttings Using Machine Vision. Trans. ASAE 33(6):2067-2073.

Sokhansanj, S. and D. M. Bruce. 1987. A Conduction Model to Predict Grain Temperatures in Grain Drying Simulation. Transactions of the ASAE 30(4):1181-1184.

Suggs, C. W., B. M. Lineberger, and S. C. Mohapatra. 1992. Automatic Feeding Transplanter. IN: Transplant Production Systems. Acta Horticulturae. 319:511-516.

Tanaka, Fujio. 1992. Direct Inserting Seeder for Culture Media. IN: Transplant Production Systems. Acta Horticulturae. 319: 551-556.

Torii, T., M. Kasiwazaki, T. Okamoto and O. Kitani. 1992. Evaluation of Graft-Take Using a Thermal Camera. IN: Transplant Production Systems. Acta Horticulturae. 319: 631-634.

Trollope, J. R. 1982. A Mathematical Model of the Threshing Process in a Conventional Combine-Thresher. Journal of Agricultural Engineering Research 27(2):119-130.

Volk, Tyler and John D. Rummel. 1987. Mass Balances for a Biological Life Support System Simulation Model. Adv. Space Res. 7(4):141-148.

Volk, Tyler and Hatice Cullingford. 1989. Crop Growth and Associated Life Support for a Lunar Farm. IN: Lunar Bases and Space Activities in the 21st Century.

Whittaker, A. Dale, G. E. Miles, O. R. Mitchell and L. D. Gaultney. 1987. Fruit Location in a Partially Occluded Image. Trans. ASAE 30(3):591-596.

N94-25372

Modelling Early Failures on Space Station *Freedom*

Final Report

NASA/ASEE Summer Faculty Fellowship Program -- 1993

Johnson Space Center

Prepared By:	Sharon E. Navard
Academic Rank:	Assistant Professor
University and Department:	Virginia Commonwealth University Department of Mathematical Sciences Richmond, Virginia 23284-2014
NASA/JSC	
Directorate:	Safety, Reliability, and Quality Assurance
Division:	Space Station *Freedom* Safety and Mission Assurance
Branch:	Analysis and Risk Assessment
JSC Colleague:	Richard P. Heydorn
Date Submitted:	August 6, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

A major problem encountered in planning for Space Station *Freedom* is the amount of maintenance that will be required. To predict the failure rates of components and systems aboard Space Station *Freedom*, the logical approach is to use data obtained from previously flown spacecraft. In order to determine the mechanisms that are driving the failures, models can be proposed, and then checked to see if they adequately fit the observed failure data obtained from a large variety of satellites. For this particular study, failure data and truncation times were available for satellites launched between 1976 and 1984; no data past 1984 was available. The study was limited to electrical subsystems and assemblies, which were studied to determine if they followed a model resulting from a mixture of exponential distributions.

INTRODUCTION

In order to accurately estimate and reduce the amount of maintenance that will be required on Space Station *Freedom*, it is necessary to understand the mechanisms that cause the failures of its components and systems. Ideally formal life tests would be conducted, where a sample of each type of component would be put on test under environmental and operational conditions identical to those under which it is to be used, and the time to failure for each would be observed. Due to time constraints, it is not always possible to observe all of the items until they fail; in this case some type of censoring mechanism is employed.

There are two basic types of censoring which have been extensively studied. Type I censoring occurs when n items are placed on test and observed for a fixed period of time t. Only the lifetimes of those which fail before time t are known; the others are said to be time censored or *truncated*. In this case the length of the test, t, is fixed, but the number of failures observed, r, is random. In Type II censoring, n items are placed on test and the test is terminated after the r^{th} item fails. In this case, the number of failures, r, is fixed in advance, but the length of time of the test, t, is a random variable. Both of these types of censoring have been treated extensively in the literature. See, for example, Barlow and Proschan (1975), Bain (1978), and Lawless (1982). For information on additional types of censoring, see McCool (1982), and Mann and Singpurwalla (1983).

Unfortunately, the situation encountered in building spacecraft is far from ideal. It is impossible to conduct meaningful life tests on earth because it is not possible to reproduce the microgravity environment. Thus to obtain data on failure of components in microgravity, it is necessary to turn to field data obtained from previously flown spacecraft. The situation is further complicated by the fact that the censoring taking place in this type of situation is neither Type I nor Type II. It is not determined in advance how many components will fail. Furthermore, the truncation times are not known in advance. The lifetime of one component may be truncated by the failure of another--for example, if the attitude control malfunctions, the satellite may fall out of orbit while all of its other components are still functioning; but none of these lifetimes are observable. Also, many components may still be operating at the last time they are observed.

The case where both the number of failures and truncation times are both random variables has not been studied nearly as extensively. In fact, there is no

consensus on what it is called. Mann and Singpurwalla (1983) and Lawless (1982) refer to this general case as random censoring, while the entry under "Random Censoring" in the *Encyclopedia of Statistical Sciences, Vol. 1* defines random censoring as a completely different situation. Nevertheless, the particular models given by Mann and Singpurwalla (1983) and Lawless (1982) are not applicable to the satellite data because they assume that failure times are identically distributed.

The approach taken in this study to determine the mechanisms that produce the failures is to propose a mechanism, and then see how well the resulting model fits what is actually observed in the data. If the theoretical survival function differs significantly from the empirical survival function, then the proposed mechanism is not what is actually producing the failures.

THE SATELLITE DATA

The failure data for over 300 satellites was compiled by Planning Research Corporation and originally published in Bloomquist, et. al. (1978), with an update in 1984. This data is currently being compiled into a data base by Loral Space Information Systems. The data base includes each component of each satellite, classified by subsystem and assembly. The times for all failures are included, and truncation times for those which did not fail.

The way the times were recorded was not consistent for all satellites. Often assemblies are turned off and on during the life of the satellite, so that they are not operating for the entire life of the satellite. Some elements, such as backup systems which are never needed, never get turned on at all. For some of the satellites, the times given were actual operating times for that assembly; for others, they were merely the time since launch, or "survival time." For the purposes of this study, operating times and survival times were treated separately.

Since the data base was still being compiled at the time of this study, it was necessary to limit the number of satellites used in order to obtain some data. Since the more recent satellites are more likely to utilize the same type of technology as Space Station *Freedom*, only satellites launched since 1976 were considered. This was a total of 28 satellites. Of these, only four had recorded operating times; the rest had recorded survival times. The data was further limited by considering only electronic assemblies, because the model considered was proposed for failure of electronic parts.

THE PROPOSED MODEL

It is generally accepted that the lifetime of electronic components has an exponential distribution, that is

$$f(t|\lambda) = \lambda e^{-\lambda t}, \quad t \geq 0,$$

where λ is a parameter determined by the failure rate. The exponential distribution possesses a unique property--it has a constant hazard rate. However, when considering many different types of electrical assemblies on different satellites, the odds are that the failure rates will not be the same for all of them. Thus the failure times do not represent a sample from a single exponential distribution, but rather from a mixture of exponentials with varying parameters. While each individual hazard rate is constant, it turns out that the hazard rate of the mixture is actually decreasing (see Barlow and Proschan (1975), p. 102).

Hecht and Hecht (1985) analyzed the original PRC data and concluded that that the failures did indeed possess a decreasing hazard rate. Heydorn, et. al. (1991) proposed a model derived from a mixture of ordinary exponentials and demonstrated that its predictions were close to the results given in Hecht and Hecht (1985). However, they did not have the original data, so their results depended on those given in the Hecht and Hecht report. This study uses the model proposed by Heydorn, et. al. and fits it to the actual satellite data.

The model is obtained by assuming that the electronic components do possess an exponential distribution with parameter λ. Thus for a fixed λ the failures are generated by a Poisson process with parameter λ. However, λ can be considered a random variable since the failure rates are not the same for all components. Taking the Bayesian approach and assuming that the prior distribution of λ is uniform on $(0, \infty)$, the posterior distribution of λ is a gamma distribution, and the resulting reliability function t is

$$R(t) = \frac{1}{\left(1 + \frac{t}{T}\right)^{\alpha+1}}, \qquad (1)$$

where α and T are parameters. The purpose of this study is to find the values of the parameters α and T which best fit the empirical survival function.

EMPIRICAL SURVIVAL FUNCTION

The next step is to estimate the empirical survival function. If all of the failure times are known, then the estimated survival function is

$$S(t) = \frac{\text{Number of observations} \geq t}{n}, \quad t \geq 0.$$

If, however, some of the survival times are unknown, then this must be modified. Kaplan and Meier (1958) introduced an estimate called the product-limit (PL) estimate which handles the case of random censoring. The estimate is

$$S_{PL}(t) = \prod_{j=1}^{k} \left(1 - \frac{\phi_j}{N-j+1}\right),$$

where k is the total number of events (including both failures and truncations) up to time t, N is the total number of events in the test, and ϕ_j is an indicator variable defined by

$$\phi_j = \begin{cases} 0 & \text{if the event is a truncation} \\ 1 & \text{if the event is a failure} \end{cases}.$$

The estimate is a product, where each term can be thought of as the conditional probability of surviving past time t_j, given survival to just prior to t_j, where t_j is the time of event j. It is a step function, which steps down at each t_j. For more information on the properties of the PL estimate, see Lawless (1982) and Peterson (1983).

The problem encountered with using the PL estimate of the reliability function is that when many of the largest times are truncations, none of the terms in the product approach zero, and hence the estimate yields only a small portion of the reliability function. For example, in the data set for satellite operating times, there were a total of 165 observations, of which 14 were failures; the rest were truncations. Furthermore, the 82 largest times were truncations. This meant that the smallest term in product was 82/83. The PL estimate of the survival function for the operating time data is:

t	$S_{PL}(t)$
50	0.9939
70	0.9878
76	0.9817
264	0.9756
1440	0.9633
3890	0.9572
4300	0.9447
10500	0.9367
14980	0.9285
20747	0.9203
29361	0.9120
34526	0.9009

The data set for satellite survival times contained 907 points, of which the last 154 were truncations. The PL estimate of the survival function for this data ranged from 0.9989 to 0.8695.

ESTIMATING THE PARAMETERS

To find the estimates of α and T in Equation (1) which give the best fit to the PL estimate of the survival function, the first approach was to use the least squares criterion. Since (1) is a nonlinear function, the nonlinear least squares routine in S-plus was used. Unfortunately, the Jacobian matrix was nearly rank deficient, for a multitude of initial values, including the estimates derived in the other approach described below. Thus another approach had to be used to estimate the parameters.

While the survival function given in (1) is nonlinear, it turns out that the inverse of the hazard function associated with it is a linear function. For a survival function S(t), the hazard function is

$$h(t) = - \frac{d\ln S(t)}{dt}.$$

The hazard function associated with the survival function given in (1) is

$$h(t) = \frac{\alpha + 1}{T + t},$$

and its inverse can be expressed as the linear function

$$\frac{1}{h(t)} = \frac{T}{\alpha + 1} + \left(\frac{1}{\alpha + 1}\right) t. \qquad (2)$$

Thus the parameters α and T can be estimated by finding the empirical hazard function, taking its inverse, and using least squares to estimate the parameters of the resulting linear equation. These coefficients can then be used to solve for α and T in (2).

The empirical hazard function was estimated by taking the log of the PL estimate of the survival function, then taking successive differences in the resulting values and dividing by the width of the corresponding time interval to approximate the derivative. A linear model was then fit to the inverse of the empirical hazard function, and the parameters α and T were determined.

RESULTS

To begin the analysis, recall that if the mixed exponential model is correct, the failures should have a decreasing hazard function. Plots of the empirical hazard function for the operating time data and survival time data are given in Figures 1 and 2, respectively. Both plots are extremely noisy, and neither one definitely indicates a decreasing hazard function. No conclusions can be drawn from these plots.

Plots of the empirical survival function and estimated theoretical survival function for the operating time and survival time data are given in Figures 3 and 4, respectively. It can be seen that neither of these demonstrates a very good fit. In particular, the empirical survival function initially drops at a much faster rate than the theoretical one for both sets of data. The data appears to display early failures, or "infant mortality" that is not adequately explained by the model. Based on these plots, the mixed exponential model does not seem to adequately explain the failures. A different model must be sought.

In conclusion, this research indicates that a mixture of exponential distributions does not adequately explain electronic failures seen in previously flown satellites. In particular, it does not model the early failures very well. A different model will be needed to explain the mechanism generating the failures.

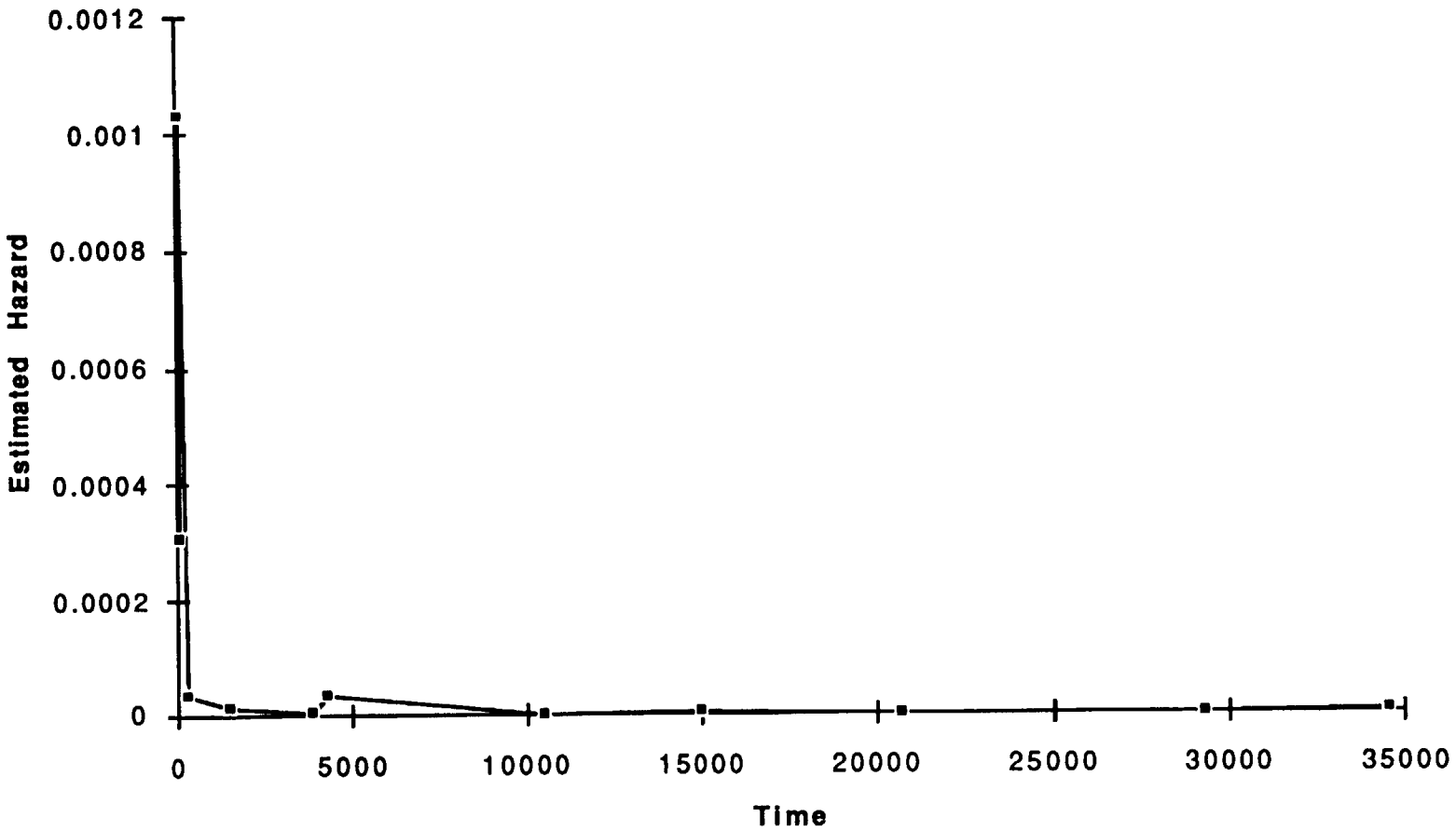


Figure 1.- Empirical hazard function of operating time data

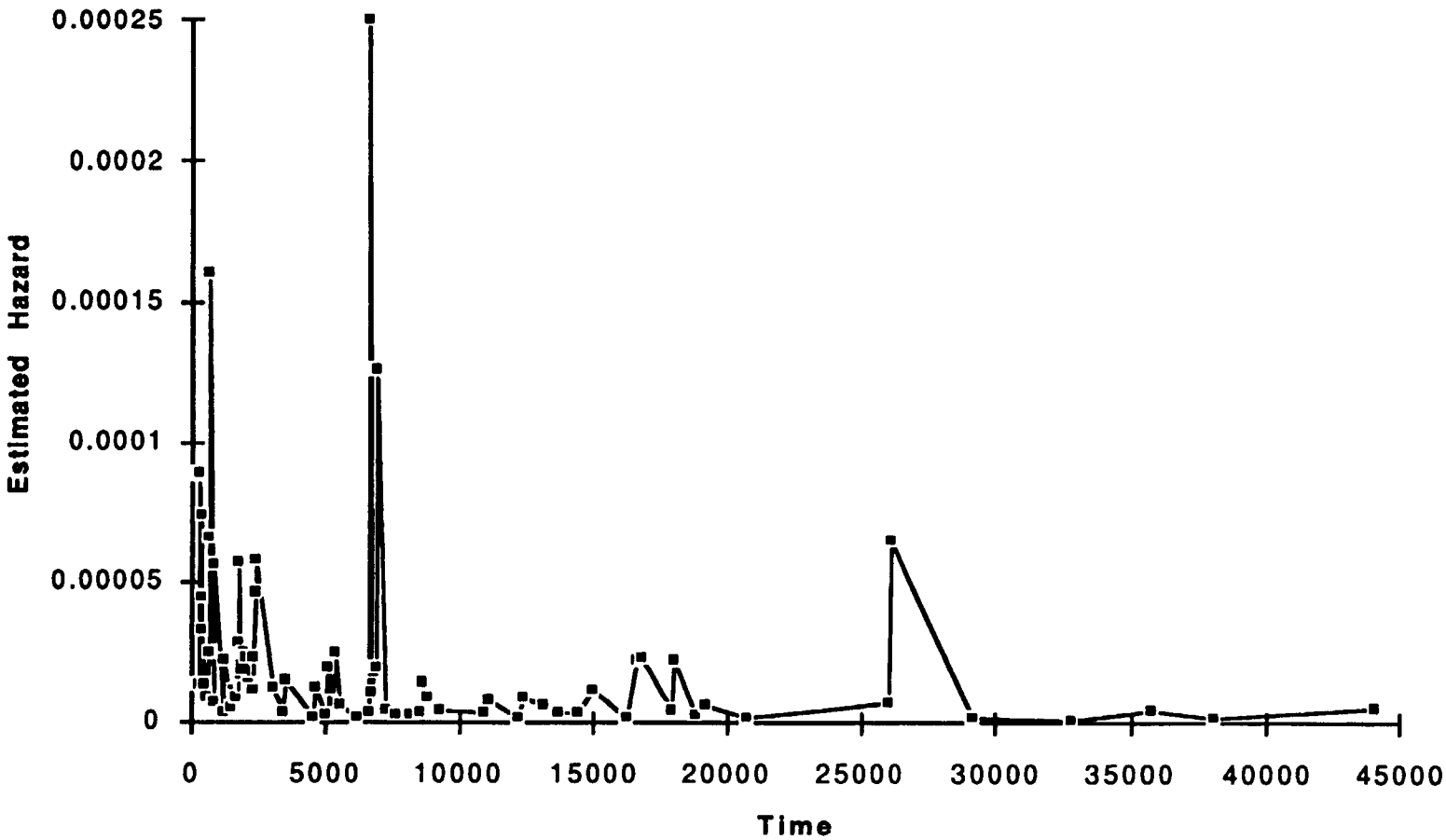


Figure 2.- Empirical hazard function of survival time data

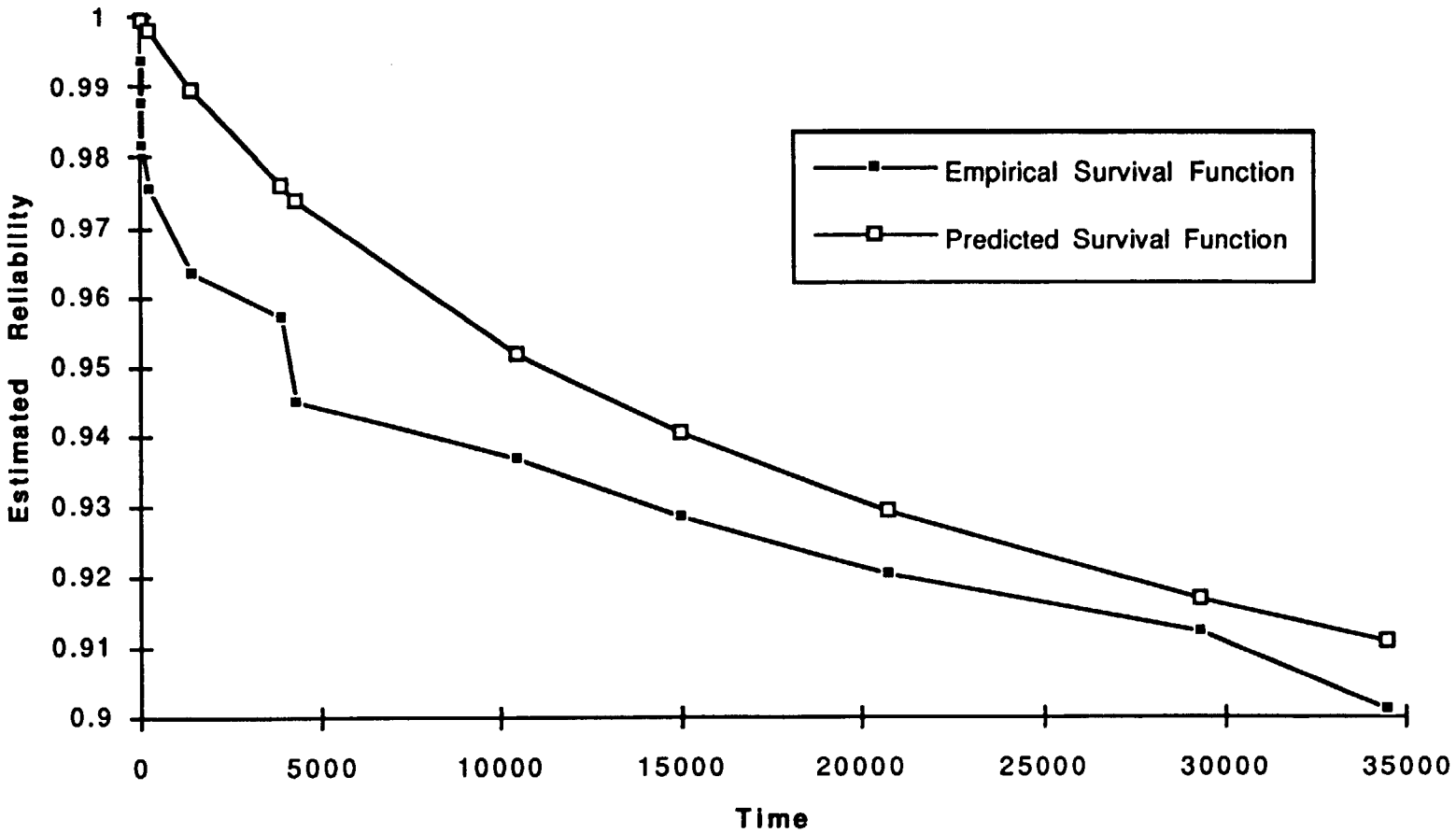


Figure 3.- Estimated survival functions of operating time data

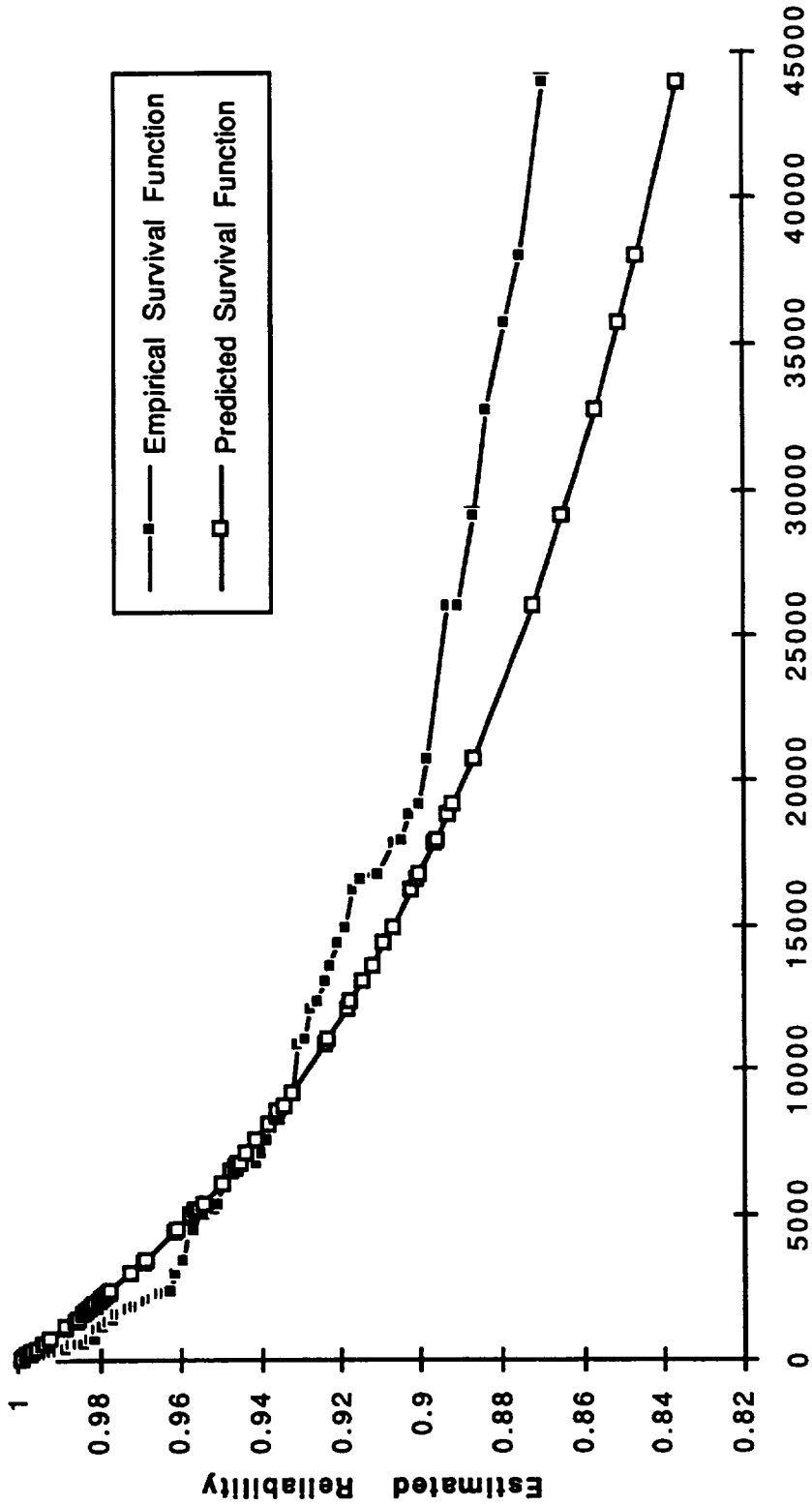


Figure 4.- Estimated survival function of survival time data

REFERENCES

Bain, L. J. (1978), *Statistical Analysis of Reliability and Life Testing: Theory and Methods*, New York: Marcel Dekker.

Barlow, R. E., and Proschan, F. (1975), *Statistical Theory of Reliability and Life Testing: Probability Models*, New York: Holt, Rinehart, and Winston.

Bloomquist, C., Anderson, V., Demars, D., Graham, W., Henruri, P., and Stiehl, G. (1978), "On-Orbit Spacecraft Reliability," PRC Report R-1863, September 30, 1978.

Hecht, J., and Hecht, M. (1985), "Reliability Prediction For Spacecraft," RADC-TR-85-229, December 1985.

Heydorn, R., Blumentritt, W., Doran, L., and Graber, R. (1991), "A Model for Projecting the Number of Early Failures on Space Station *Freedom*," NASA/JSC preliminary report.

Kaplan, E. L., and Meier, P. (1958), "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association*, **53**, 457-481.

Lawless, J. F. (1982), *Statistical Models and Methods for Lifetime Data*, New York: John Wiley & Sons.

McCool, J. I. (1982), "Censored Data," in *Encyclopedia of Statistical Sciences Vol. 1*, eds.S. Kotz and N. L. Johnson, New York: John Wiley & Sons, PP. 389-396.

Mann, N. R., and Singpurwalla, N. D. (1983), "Life Testing," in *Encyclopedia of Statistical Sciences Vol. 4*, eds.S. Kotz and N. L. Johnson, New York: John Wiley & Sons, PP. 632-639.

Peterson, A. V. (1983), "Kaplan-Meier Estimator," in *Encyclopedia of Statistical Sciences Vol. 4*, eds.S. Kotz and N. L. Johnson, New York: John Wiley & Sons, PP. 346-351.

N94-25373

A COMPARATIVE EVALUATION PLAN FOR THE MAINTENANCE, INVENTORY, AND LOGISTICS PLANNING (MILP) SYSTEM HUMAN-COMPUTER INTERFACE (HCI)

Final Report
NASA/ASEE Summer Faculty Fellowship Program - 1993
Johnson Space Center

Prepared By:	Scott P. Overmyer
Academic Rank:	Assistant Professor
University and Department:	Minot State University College of Business Minot, North Dakota 58707
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Flight Crew Support
Branch:	Crew Station
JSC Colleague:	Frances E. Mount
Date Submitted:	August 4, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

The primary goal of this project was to develop a tailored and effective approach to the design and evaluation of the human-computer interface (HCI) to the Maintenance, Inventory and Logistics Planning (MILP) System in support of the Mission Operations Directorate (MOD). An additional task that was undertaken was to assist in the review of Ground Displays for Space Station Freedom (SSF) by attending the Ground Displays Interface Group (GDIG), and commenting on the preliminary design for these displays.

Based upon data gathered over the 10 week period, this project has hypothesized that the proper HCI concept for navigating through maintenance databases for large space vehicles is one based upon a spatial, direct manipulation approach. This dialogue style can be then coupled with a traditional text-based DBMS, after the user has determined the general nature and location of the information needed. This conclusion is in contrast with the currently planned HCI for MILP which uses a traditional form-fill-in dialogue style for all data access and retrieval.

In order to resolve this difference in HCI and dialogue styles, it is recommended that a comparative evaluation be performed which combines the use of both subjective and objective metrics to determine the optimal (performance-wise) and preferred approach for end users. The proposed plan has been outlined in the previous paragraphs and is available in its entirety in the Technical Report associated with this project. Further, it is suggested that several of the more useful features of the Maintenance Operations Management System (MOMS), especially those developed by the end-users, be incorporated into MILP to save development time and money.

INTRODUCTION

The primary goal of this project was to develop a tailored and effective approach to the design and evaluation of the human-computer interface (HCI) to the Maintenance, Inventory and Logistics Planning (MILP) System in support of the Mission Operations Directorate (MOD). An additional task that was undertaken was to assist in the review of Ground Displays for Space Station Freedom (SSF) by attending the Ground Displays Group (GDIG), and commenting on the preliminary design for these displays.

In order for the MILP and other ground and space-based systems to be effective, it is imperative that the HCI is convenient and easy to use so that user personnel can spend time solving problems, instead of grappling with the user interface.

The MILP project was divided into several tasks as listed below, which were completed over a period of 10 weeks. Tasks "a" through "e" were essentially background tasks for production of an evaluation plan for the MILP HCI.

(a) Review User and Task Data for MILP
(b) Develop Scenario of Operations
(c) Develop Storyboard of the Scenario
(d) Conduct Storyboard/HCI Walkthrough with Users
(e) Build Interactive Rapid Demonstration Prototype
(f) Develop Full Comparative Evaluation Plan

REVIEW USER AND TASK DATA FOR MILP

The first task was to gain a sufficient understanding of the potential users of MILP and the tasks that they perform both with and without using automated systems. This was accomplished by: (1) Reading documentation about both MILP and similar systems, (2) observing the current MILP HCI in use, (3) observing In-Flight Maintenance (IFM) personnel using the Maintenance Operations Management System (MOMS) during a shuttle mission, and (4) by interviewing management personnel about their goals for the MILP system. More detailed information can be found in an accompanying Technical Report which is located in the Human-Computer Interaction Laboratory at JSC.

The MILP system is intended to support, as its title implies, maintenance, inventory and logistics planning for future space missions including Space Station Freedom (SSF). The functional requirements for MILP have been divided into 5 task areas which are as follows:

1) Use and Augment Support Data and Documentation

The purpose of this function will be to collect data about space craft and space craft systems and subsystems and make it available for use by support crew on

the ground. This function also will provide the capability to augment textual information with locally constructed diagrams, video, text, etc. Generally speaking, this data is used to help ground personnel solve problems which may arise onboard the space craft. This data and associated maintenance procedures may then be uplinked to on-orbit crews upon request.

2) Maintain On-Orbit Replaceable Unit (ORU) Material History Data

The purpose of the Maintain On-Orbit ORU Material History Data will be to allow the user to update the initial history data for each ORU to reflect current status, usage, and other significant life cycle events of the ORU's.

3) Manage Onboard Inventory

The Manage Onboard Inventory function will provide the user with the capability to plan and execute inventory/stowage and resupply/return operations. The user will be able to: track the current status and configuration of inventory and stowage; use nominal procedures, crew reports, and the resupply/return manifest to update the current onboard inventory database; supply the location of stowage and the status of inventory to the program; use the program to help integrate the resupply/return manifest requirements and to provide to Level II for manifest development; develop onboard inventory databases for future increments; and store onboard inventory data from past increments.

4) Compile Resupply Return Requirements

This function will allow the user to combine resupply return requirements obtained from TCATS workstations. Inputs to the process will be: delta , WP, user, and RUPSM resupply return requirements. The system will be capable of receiving core systems resupply return requirements and combining them sorted by user query. In addition, the user will be able to transfer the combined resupply return requirements to the LIS for analysis.

5) Maintain Physical Configuration

The Maintain Physical Configuration function will allow the user to view a hierarchical representation of historical, current, and planned station physical configurations. Using this function, the user will be able to engage in "what-if" scenarios regarding station configuration. With this function, the user will be able to store, access and edit (if authorized) the last 2 increments and 10 future increments of station physical configuration data, and will be able to access an archive of data older than 2 increments. MILP will also maintain a time-tagged log of all changes to the current station physical configuration database. The user will also have a MILP tool which allows the hierarchical modeling of historical, current, and planned station physical configurations.

An information system currently exists to support procedure development and other onboard maintenance operations related to Space Shuttle missions. This system was also studied for applicability to the MILP project. The Maintenance Operations Management System (MOMS) is a computer-based system designed to support the development of maintenance procedures for uplink to the crew. The MOMS system supports some similar sub-tasks as described above in MILP Task 1 - Use and Augment Support Data and Documentation, however, has a much small scope than this task in the MILP system. MOMS does not provide support for the other 4 tasks required by MILP, however, it is useful to look at MOMS since it provides support for the only near-real-time tasking performed by a console operator using either MOMS or MILP.

MOMS consists of a SUN 3/260 workstation equipped with a Parallax video processing board, one monochrome and one color monitor, a keyboard, and an optical mouse. Peripheral to the system is a suite of video equipment consisting of 2 optical disk storage units, a video tape player/recorder, video interface devices, an additional color monitor for live video signal display, and a still frame camera for capturing and displaying images of objects, manuals and photographs. Software consists of the InterLeaf publishing system, and the MOMS custom software for video image processing.

MOMS includes the Interleaf desktop publishing software which allows users to create generic documents and incorporate text, graphics, tables and figures into the document. These documents are the same as any standard desktop publishing system can produce, however, the MOMS users (IFM) have developed some standard documents specifically for creating paper procedures and flight notes for uplink during flights, and log pages for use recording operator's notes during console shifts. These documents have been put on the Interleaf menus for operator use. An illustration of the Interleaf screen is shown in Figure 1.

The MOMS users have also developed some customized components for use with the custom procedure documents. These components are items which can be inserted into the document at any point by placing the cursor in the document, and then selecting the desired insert from a menu.

The IFM users keep a file of checklists which have been developed in response to various problems, and which may be recalled for use or modification at a later date. This file is called the Supplemental Checklist File and is stored on disk in a "cabinet". This cabinet is divided into several "file drawers" alphabetically, according to the title of the procedure.

Based upon observations of two MOMS users during a mission, it appeared that the primary usage for MOMS was and is procedure development, construction, and publishing. This activity mainly involves the text/publishing portion of the system which is displayed on the monochrome display screen. Occasionally, a

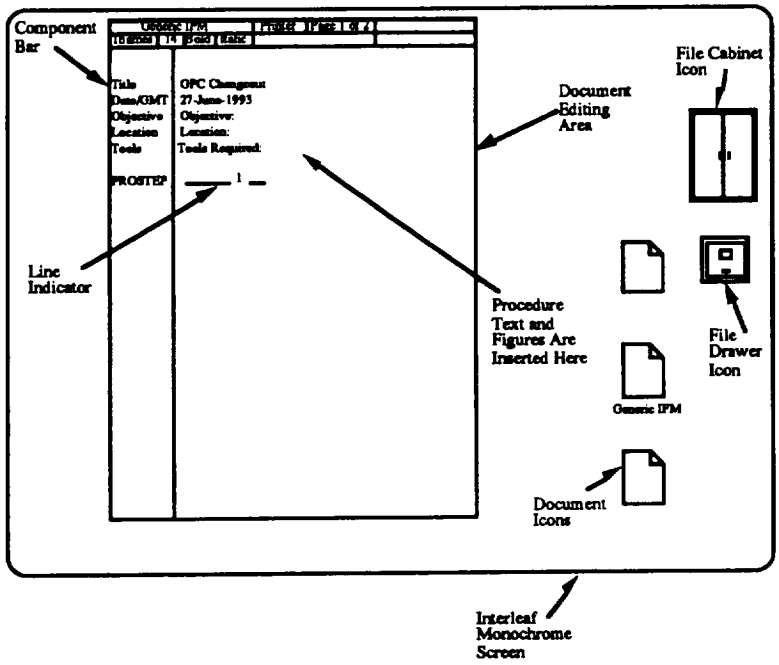


Figure 1 - Interleaf Screen with Open "Generic IFM" Procedure Document

user would reference a photograph stored on one of the two video disk systems via the MOMS video interface, but more often the user would turn to the large notebooks of photographs stored in a nearby cabinet. A knowledgeable user stated that the search time for information from photographs in the books was generally faster than with the system. In addition, he stated that the resolution of the photos on the screen was lower than that in the books making it more difficult to see very small objects displayed on the screen. This is, in part, supported by the fact that NTSC video resolution is usually not greater than 500 horizontal lines, while photos are generally 2000 lines; however, the clarity of the photographic images displayed on the color CRT appeared quite good for many depicted objects.

One of the MOMS features most relevant to MILP, the ability to capture still video and include it in procedures, was not often used since the technology at the time MOMS was built did not allow color-to-monochrome conversion at sufficent resolution.

Some specific observations related to the operation of MOMS should be noted for input to the MILP design process:

(1) Custom features designed by the users to facilitate procedure development (e.g., custom components, procedure forms (which include automatic renumbering of procedure steps), other user-developed forms, and special clip

art drawings, should be included in the basic MILP design. If Interleaf is to be used, these features could be directly transferred to the new system.

(2) Image capture and translation algorithms which provide acceptable color image capture and color-to-grayscale translation should be employed to allow user to view color images, and include them in procedure documents especially when the final product is in black and white.

(3) All equipment should be kept in perfect working order, and initial and periodic training should be given as to its proper use. This training (and the design of the user interface) should be oriented toward occasional users, since maintenance anomalies and changes to SSF configuration occur infrequently. Since SSF will be monitored continuously after insertion, use of MILP will be more frequent than the mission-oriented use of MOMS.

(4) A method should be devised for characterizing and storing information about photographs which allows for the searching of image features not originally intended to be referenced. Users often search for pictures and locations of component parts, the existence of which is often not indicated in the title or the indices of the photograph. Search of photographs by title alone is insufficient to make the photo database useful to MOMS and MILP users.

DEVELOP SCENARIO OF OPERATIONS

The following scenario represents the actions an operator would be required to perform using MILP and other systems to respond to a specific Maintenance Contingency. This scenario contains operator and other physical and communication actions which are performed without the use of MILP. These actions are included in the scenario for completeness.

The core upon which this scenario is built is taken from Scenario: IPS-MSN-16 MILP: Maintenance/Contingency Support; JSC-13601; SSFP Integrated Planning System - Project Plan, Volume 3, Ops Concept - Appendix C. Within this core scenario, there are 3 possibilities, the most interesting of which is Scenario "c" - "Development of a new maintenance procedure and its associated activity definition form". Since this document contains an inter-related collection of scenarios which reference each other, IPS-MSN-16 serves as the base scenario, with portions of others included to complete the entire scenario to be implemented in the rapid prototype. This scenario covers the MILP Task Area #1, "Use and Augment Support Data and Documentation". The particular maintenance contingency which has been selected is based upon data collected from a variety of sources including:

- Published IFM procedures (primarily GPC Replacement - Multi GPC)
- "Ops Concept" scenarios
- Interviews with MOMS and potential MILP users
- Other documents

SCENARIO - During a payload rack installation procedure, a metal object that crew member was carrying accidentally slipped from the crew member's grip. As the crew member was turning at a fairly high rate of speed, the object entered an open computer cabinet and pierced a wiring harness. The systems operator received an alarm at the console indicating that inter-computer communication had ceased. Closer inspection by the crew determined that several wires in the harness had been severed. The accident has occurred inside a pressurized module.

The systems operator has contacted the Ops Support Officer (OSO) to discuss maintenance options. The unlikelihood of such an occurrence precluded the stocking of a replacement wiring harness. Together, the system operator, the OSO, and the Flight Director have decided that the maintenance action must be taken prior to the crew's departure from SSF. The OSO is a MILP end-user.

1) The OSO informs the Ops Planner that the Short Term Plan (STP) will be impacted (via voice).

2) The OSO (or possibly the OSO Support Personnel or Inventory and Stowage Officer) using MILP, checks the inventory and determines that the spare wiring harness required to correct the problem is not onboard.

3) The OSO and the systems operator explore viable maintenance options such as routing signals to another unit, interchanging units, or bypassing the failed unit. It is determined that these options are unavailable due to lack of redundancy and backups for this particular system.

4) The OSO reviews the data currently in MILP to support the maintenance options analysis, including photographs of the rear of the computer equipment in place showing the location and accessibility of the damaged component. The OSO also finds some video of the crew replacing a computer unit on a previous flight. This video shows the removal and replacement of the wiring harness in question.

5) For data that is not available in IPS, the OSO logs on to EDLS to browse for an engineering drawing showing the wiring harness in detail, and brings a copy of the drawing into MILP.

5a. From an IPS workstation, an authorized EDLS user opens a window to log onto EDLS, browses the data, and selects products for retrieval. Alternatively, the OSO may contact someone else who is custodian of engineering drawings, and request the drawings be retrieved.

5b. The IPS user transmits a file transfer request from the IPS workstation via the FTP file transfer service as specified in the ICD.

5c. The transfer request is processed and the file is transferred from EDLS directly back to the requester's (MILP user's) personal data area.

6) After discussions and consultations with other Subject Matter Experts (SME's), the OSO then assembles the data into 2 option related packages. For this scenario, these options are: a) to remove the harness and repair the severed wires, or b) to replace the harness with another harness using the connectors from the damaged harness.

7) As the data is assembled/manipulated by the OSO, on user's request a list of specific data items being analyzed at the time of request (i.e., "a snapshot") is automatically captured by MILP to provide a record of actions taken by the OSO. The records provide documentation of the analysis trail leading to the solution and are stored in the MILP personal data store until the maintenance action is completed. This data is then archived.

8. The OSO and system operator(s) view the data in the option packages simultaneously to permit a complete analysis of the options. Upon completion of the analysis, the OSO and system operator(s) select option "a", to repair the severed wires on the damaged harness, and obtain approval of the selection by the Flight Director.

9. The OS informs the Systems Operations Data File (SODF) Increment Coordinator that a new maintenance procedure is required (via voice).

10. The OSO builds the maintenance SOP (Using Interleaf) from the assembled data and coordinates with the SODF Increment Coordinator for the SOP's import into PDAC. The development and validation of the flight procedure takes place using PDAC. The OSO has the capability to personally perform this task using PDAC.

11. The OSO fills out an Activity Definition Form (ADF) and contacts the Ops Planner (via voice) to inform him/her that the activity is ready for incorporation in the Short Term Plan. The OSO saves the ADF in the MILP user's personal data store.

11a. The OSO accesses the Activity Definition Interface (ADI) which provides an Activity Definition Form (ADF) to be filled out.

11b. The OSO fills out the ADF. The ADI provides prompts to solicit the correct information for the activity definition. The OSO may enter just the basic information (i.e., activity name, duration, and window), and the Consolidated Planning System (CPS) user (the Ops Planner), through an iterative process with the OSO, will complete the other required information. Or the OSO may complete the ADF through an iterative process, making it available to the CPS user when complete.

11c. The OSO/CPS user store the resulting ADF in the shared data store. The OSO informs a CPS user that the ADF is stored. The OSO saves a copy of the ADF in his own database.

11d. The CPS user accesses and reviews the newly-input ADF, and in an iterative process, consults with the OSO and other SME's as required to insure the ADF(s) is as complete and accurate as possible. Data will also be obtained from other IPS subsystems to complete the ADF(s).

11e. After ADF review, the CPS user imports the ADF and creates the Activity Definition. (The Activity Definition will require concurrence from both the OSO and the CPS user.) *Non-MILP Activity*

11f. The activity definition is verified and promoted by the CPS user to the Master Data Store for input into the planning process. The definition could later be stored as a "Master Activity Definition" in the Master Data Store if approved by the Ops Plan team. *Non-OSO/MILP Activity*

11g. The CPS user uses the activity definition as building blocks in the activity timeline development. *Non-OSO/MILP Activity*

12. Upon approval by the flight director, the procedure is prepared for uplink to the crew and the OSTP is updated and uplinked. In the Post-MTC timeframe, an OSTP will be developed by extracting the onboard portion of the STP and creating an onboard version within the CPS. *Non-OSO/MILP Activity*

13. If the above activity crosses shifts or is interrupted for other reasons the OSO initiates MILP action to save the records providing documentation of the analysis trail so it can be retrieved for work on the next shift or at some later time.

14. The procedure is reviewed with the crew.

15. The OSO monitors the crew execution of the procedure and responds as needed to queries. If deviations from the procedure occur, the OSO annotates a copy of the flight procedure and logs material history and/or physical configuration changes.

16. Upon completion of the maintenance action, the OSO updates the hardware's material history using the Configuration and Verification Reporting System (CVRS), and sends a copy of the updated material history to the appropriate Engineering Support Center (ESC) (via FTP or Fax).

17. The OSO updates the MILP database to document changes to the vehicle's configuration as a result of the maintenance activity and transmits via TMIS a copy of the configuration changes to the SSFP Systems Engineering and Integration (SE&I) organization (or Configuration Management, TBD).

DEVELOP STORYBOARD OF THE SCENARIO

The object of this task was to instantiate the operational scenario in a tangible HCI design which could be evaluated by project stakeholders, particularly end-users. The storyboards were paper-based, consisting of the types of illustrations shown in Figure 2. This storyboard was used to describe the high-level HCI concept. The complete storyboard is included in the associated Technical Report for this project.

This series of storyboarded views of the MILP HCI made it possible to perform a critical analysis of the MILP operational scenario prior to the HCI walkthrough. Some significant modifications to the scenarios were made and reflected in the storyboard.

First, scenario steps 5 through 5c were shown as automated in the storyboard. In the original scenario, the MILP user is required to open another window and manually log-on to another system to search a database for relevant information. This involves searching heterogeneous databases using different searching procedures and different HCI designs to accomplish a single activity. As an alternative, the storyboards show the dialogue design as one in which a user requests information about an item of interest, and is presented with a menu of potentially available data types. The user selects one or more of the data types (e.g., photos, drawings, video, procedures, etc.) and the MILP system automatically logs on to the system on which each data type resides and retrieves a list of available data items. The user then selects the items desired,

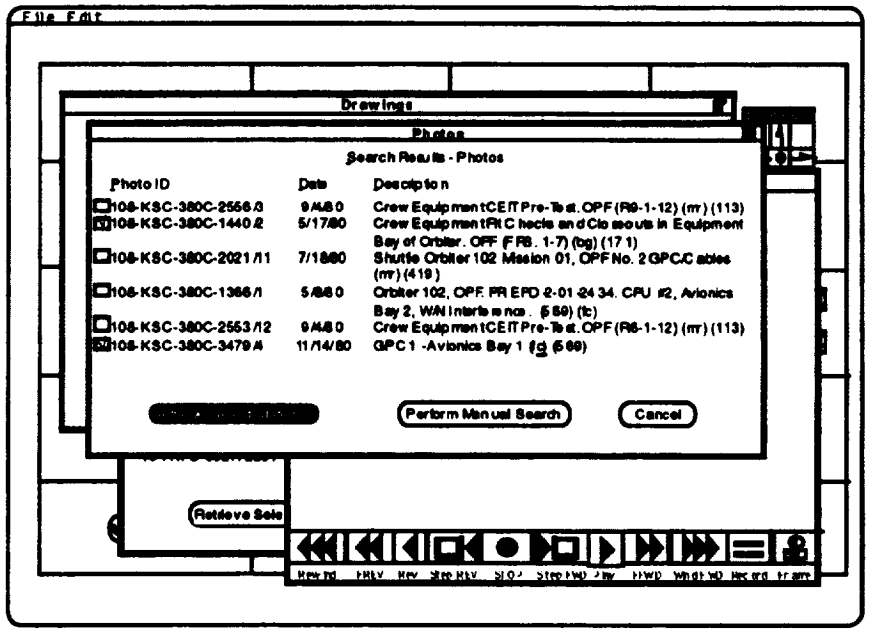


Figure 2 - Photo Database Search Results

and MILP retrieves them from the remote system into the user's workstation (with the exception of video which is handled differently). Figure 2 is an example of the storyboard for photographic database search results. This change does not preclude the possibility for the user to log on to a remote system and perform a manual search, however, it automates the process for all but the most complex searches for maintenance data items.

CONDUCT STORYBOARD/HCI WALKTHROUGH WITH USERS

A walkthrough of the scenario, storyboards, and the HCI concept was held on July 14, 1993 in the HCIL. In attendance were representatives of MOD end users of MOMS and MILP, the principal investigator, and another scientist from JSC HCIL. The purpose of the walkthrough was to validate the operational scenario, the storyboard, and the HCI concept for MILP.

The most useful information was collected during the scenario walkthrough, since the scenario was in sufficient detail for the users to offer specific comments. User comments were recorded and the scenario was modified accordingly. Some of the more substantive modifications were:

(1) The activity causing the accidental damage to SSF equipment was changed to a "payload rack installation procedure" to more closely reflect a real incident possibility.

(2) The video access method was defined to reflect remote access to analog video tapes and equipment, instead of digitally stored video. This change required a modification of the HCI prior to prototype implementation.

(3) It was determined that the Activity Definition Form (ADF) was undefined at this time, and there was some debate about its final composition and access. As a result, it was recommended that the ADF be reflected in the form of an Interleaf-based procedure at this time.

(4) Checkpointing of MILP activity for archival purposes was a fuzzy concept and currently not well defined. This provided the opportunity to develop a new concept and reflect it in the prototype.

Following the scenario walkthrough, the storyboard was reviewed. With few small exceptions, the HCI concept met the approval of the users.

BUILD INTERACTIVE RAPID DEMONSTRATION PROTOTYPE

While the storyboard can provide guidance for HCI designers, the best that can be expected is to transmit the concept and, in part, the "look" but not the "feel" of the HCI to the users. Also the storyboards are not in sufficient detail to be useful to evaluation the HCI in a formal sense.

The rapid demonstration prototype was selected as the technique used to complete the definition of the HCI concept and specification of the details of operation of the HCI for the MILP system. The rapid prototype was built primarily as a vehicle for expressing the trial benchmark task to be incorporated into the quantitative portion of the evaluation plan for the MILP HCI.

The prototype was built on a Macintosh Quadra 950, with a high-resolution 19" color monitor, extended keyboard, and mouse cursor control device. The software used for constructing the prototype skeleton was Supercard v1.6. Graphics, photos, and drawings were prepared using MacDraw Pro and Adobe Photoshop. Video was digitized from VHS-video into Quicktime digital format for insertion into the prototype. With the exception of HCI detailed information, the rapid demonstration prototype followed the scenario and the storyboard outline. Those dialogue types originally portrayed in the storyboard were subsequently implemented in the prototype.

The prototype is available as a stand-alone application. It is intended to be used to determine a benchmark for the comparative evaluation of human performance while users are employing alternative HCI designs to MILP.

DEVELOP FULL COMPARATIVE EVALUATION PLAN

The previous tasks were in preparation for the development of this comparative evaluation plan for the MILP HCI. Until the purpose and operational methods of the system are known, it is impossible to determine how to properly evaluate the suitability of the HCI design for the intended system mission.

The following plan consists of 2 components; subjective and objective. The evaluation starts with the objective assessment of the MILP HCI, collecting data on a number of metrics related to user performance during maintenance scenarios. After each user trial, the HCI is subjectively assessed by the user, while completing a usability questionnaire.

The objective component is comprised of a series of benchmark task comparisons across all functional areas in MILP. In order to provide a structured evaluation procedure, a Usability Specification Table [WHIT88] is constructed. An example of several possible entries in the table is shown in Table 1. The measurement concept for each attribute of interest in the Usability Specification Table is based on a portion of a scenario like the one described in a previous section. Scenarios which include activities, tasks, and subtasks are required which are representative of user activities across all MILP Task Areas. The most critical type of scenario (i.e., maintenance contingency) has already been constructed during this project, and is ready to use in the evaluation. At least 4 more detailed scenarios are required which cover MILP task areas: Maintain On-Orbit Replaceable Unit Material History Data, Manage Onboard Inventory, Compile Resupply Return Requirements, and Maintain Physical Configuration.

Table 1 - Example Entries From a MILP Usability Specification Table

ATTRIBUTE	MEASURING CONCEPT	MEASURING METHOD	WORST CASE	PLANNED LEVEL	BEST CASE	NOW LEVEL
Usability	Maintenance Contingency Task	Time to Complete the Scenario	1 Hour	30 minutes	20 minutes	N/A
Usability	Maintenance Contingency Task	Time spent in errors	15 minutes	2 minutes	0 minutes	N/A
User's Evaluation	Question-naire Score	Semantic Differential Score	7 (strongly negative)	3 (somewhat positive)	1 (strongly positive)	N/A
Preference Over Existing MILP HCI	Question-naire Score	Semantic Differential Score	Same as Existing		None Prefer Existing	N/A

As suggested by Lewis (1992) and Chin et al (1988), a subjective assessment instrument should be administered after the users have executed the selected scenarios using the HCI under evaluation. For MILP, it is felt that the usability questionnaire which was constructed and validated by Lewis (1992) is most suitable for the subjective evaluation. The questionnaire is shown in the complete technical report.

CONCLUSION

In conclusion, based upon data gathered over the 10 week period, this project has hypothesized that the proper HCI concept for navigating through maintenance databases for large space vehicles is one based upon a spatial, direct manipulation approach. This dialogue style can be then coupled with a traditional text-based DBMS, after the user has determined the general nature and location of the information needed. This conclusion is in contrast with the currently planned HCI for MILP which uses a traditional form-fill-in dialogue style for all data access and retrieval.

In order to resolve this difference in HCI and dialogue styles, it is recommended that a comparative evaluation be performed which combines the use of both subjective and objective metrics to determine the optimal (performance-wise) and preferred approach for end users. The proposed plan has been outlined in the previous paragraphs and is available in its entirety in the Technical Report associated with this project. Further, it is suggested that several of the more useful features of the MOMS system, especially those developed by the end-users, be incorporated into MILP to save development time and money.

REFERENCES

CHIN88 — Chin, J.P, Diehl, V.A., and Norman, K.L. (1988) Development of an instrument measuring user satisfaction of the human-computer interface. In Proceedings of the ACM CHI'88 Conference on Human Factors in Computing Systems, 213-218.

JSC-13193 — Space Station Control Center User Detailed Functional Requirements, April, 1990.

JSC-13347 — Space Station Control Center Maintenance, Inventory, and Logistics Planning Subsystem Functional Requirements, February, 1992.

JSC-13407 — Space Station Control Center Maintenance, Inventory, and Logistics Planning Subsystem Specification, March, 1992

JSC-13601 — SSFP Integrated Planning System - Project Plan, Volume 3, Ops Concept - Appendix C

LEWI92 — Lewis, J.R. (1992) Psychometric evaluation of the post-study system usability questionnaire: The PSSUQ. In Proceedings of the Human Factors Society 36th Annual Meeting, 1259-1263.

MIL-STD-1472 — Military Standard, Human Engineering Design Criteria for Military Systems, Equipment, and Facilities, Revision D, April, 1991.

NASA-STD-3000 — Man-Systems Integration Standards, Volume IV

SSP-30540 — Space Station Freedom Program Office Human Computer Interface Guide, June, 1991.

SSP-30570 — Space Station Freedom Program (SSFP) Flight Human-Computer Interface Standards.

WHIT88 — Whiteside, J., Bennett, J., and Holtzblatt, K. (1988) Usability engineering: Our experience and evolution. In M. Helander (ed) Handbook of Human-Computer Interaction. Amsterdam: North-Holland.

N94-25374

MODELING AND ANALYSIS OF SELECTED SPACE STATION COMMUNICATIONS AND TRACKING SUBSYSTEMS

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Elmer Raydean Richmond
Academic Rank:	Associate Professor
University & Department:	Tarrant County Junior College South Campus Computer Science Department Fort Worth, Texas 76119
NASA/JSC	
Directorate:	Engineering Directorate
Division:	Tracking and Communications Division
Branch:	Systems Engineering Branch
JSC Colleagues:	David Overland Sally Stokes

Date Submitted:	August 6, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

The Communications and Tracking System on board Space Station Freedom (SSF) provides space-to-ground, space-to-space, audio, and video communications, as well as tracking data reception and processing services. Each major category of service is provided by a communications subsystem which is controlled and monitored by software. Among these subsystems, the Assembly/Contingency Subsystem (ACS) and the Space-to-Ground Subsystem (SGS) provide communications with the ground via the Tracking and Data Relay Satellite (TDRS) System. The ACS is effectively SSF's command link, while the SGS is primarily intended as the data link for SSF payloads.

The research activities of this project focused on the ACS and SGS antenna management algorithms identified in the Flight System Software Requirements (FSSR) documentation, including:

- Software modeling and evaluation of antenna management (positioning) algorithms.
- Analysis and investigation of selected variables and parameters of these antenna management algorithms, i.e., descriptions and definitions of ranges, scopes, and dimensions.

In a related activity, to assist those responsible for monitoring the development of this flight system software, a brief summary of software metrics concepts, terms, measures, and uses was prepared.

The results of these research activities can be summarized as follows:

1) A software prototype of the ACS antenna control algorithms was developed to capture the basic positioning requirements of the ACS. This model was exercised repeatedly with input files of representative (expected) TDRS data and random TDRS data. The bias and translation algorithms for converting TDRS data to ACS coordinates appear to be correct, although considerable uncertainty is created by inconsistent and/or erroneous data values, ranges, and dimensions in the FSSR; moreover, not all such uncertainties have been resolved.

2) A very brief exploration of the SGS antenna management (positioning) algorithms was completed. SGS-peculiar functions were encoded and combined with basic ACS-related functions to allow a quick look at the SGS. This model was exercised using a limited sample of input files. At first look, the algorithms appear correct, but the SGS system description suffers from inconsistencies and uncertainties similar to those found for the ACS.

3) A considerable amount of the research period was devoted to investigation, discussion, and resolution of characteristics of selected ACS or SGS variables and parameters.

4) A brief written introduction to software metrics was prepared.

INTRODUCTION

The Communications and Tracking System (C&TS) planned for Space Station Freedom (SSF) will provide space-to-ground, space-to-space, audio and video communications, as well as tracking data reception and processing and other services necessary to support the space station mission. [1] [2] Each major category of service, identified in Figure 1, is provided by a dedicated C&TS subsystem.

Among these C&TS subsystems, the Assembly/Contingency Subsystem (ACS) provides space-to-ground communications via S-band Single Access Tracking and Data Relay Satellite System (TDRSS) services; Figure 2 identifies the major ACS hardware components. The ACS interfaces with on-board audio and on-board data systems and carries audio and/or data signals. The ACS is capable of data rates up to 192 kbps in the high data rate mode, the normal mode of operations planned; the actual rate is determined by the mix of audio versus data present in the signal. In the low data rate mode, a rate of 12 kbps is possible. There are two identical sets of ACS hardware elements planned as shown in Figure 2, with software control of which particular set (called a 'string') is in service; included in these redundant hardware sets are steerable high rate S-band horn antennas. The low data rate mode will use an omni-directional antenna and is essentially a backup system to the steerable antennas. The primary function of the ACS is to provide the command link between ground stations and the SSF.

The Space-to-Ground (SGS) subsystem also provides space-to-ground return link communications via Ku-band Single Access TDRSS services; Figure 3 identifies its major hardware components. The SGS interfaces with on-board audio, on-board video, and payload data systems. The primary function of the SGS is to provide the data link between ground stations and the SSF payloads.

This project focused on software modeling and analysis of the ACS and the SGS subsystems, with emphasis being given the ACS due to its critical role as the SSF's command link. There are, however, strong similarities between the ACS and SGS, such that the findings related to the ACS have general applicability to the SGS as well. The basic purpose of the research was to explore, via a software modeling frame of reference, the ACS and SGS subsystems as currently described by the C&TS Flight System Software Requirements (FSSR) documents. [3] [4] This exploration, in turn, had as its goals the identification, description, and, where possible, resolution of inconsistencies or anomalies found in the FSSR descriptions of these two subsystems and their requirements. The research was further limited to evaluation of the ACS High Gain Antenna Management and the SGS Antenna Management subcapabilities which deal basically with antenna positioning functions to maintain communications via the TDRSS. The specific research activities included:

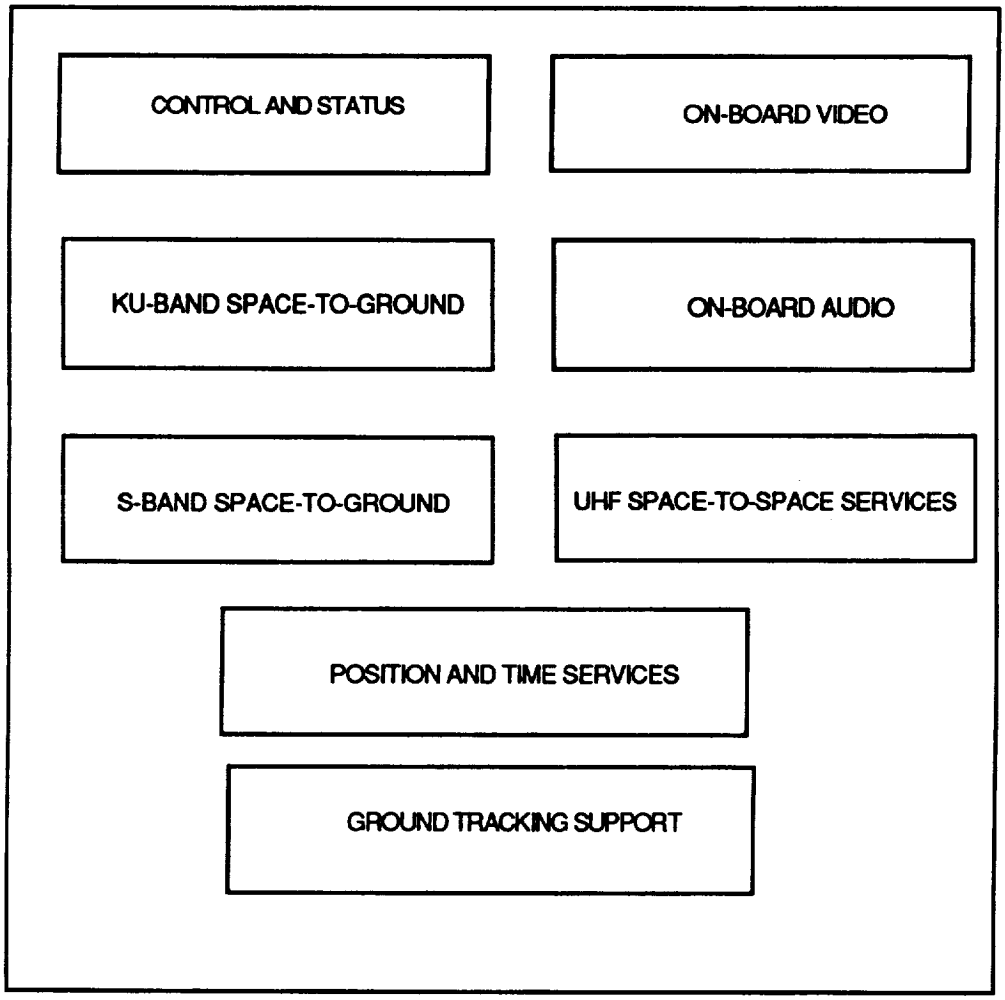


Figure 1.- Communications and Tracking Functions

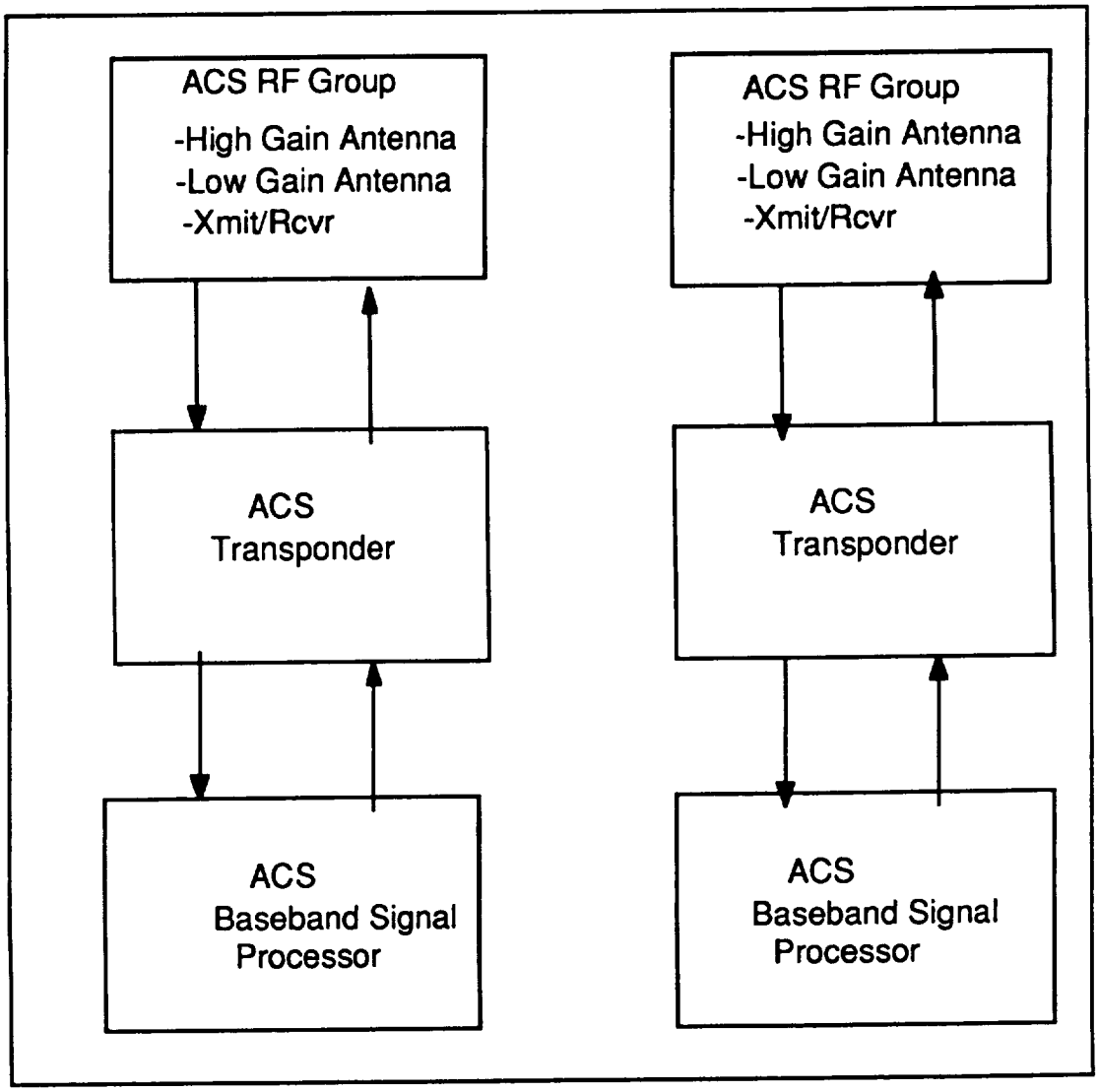


Figure 2.- Assembly/Contingency Subsystem Hardware

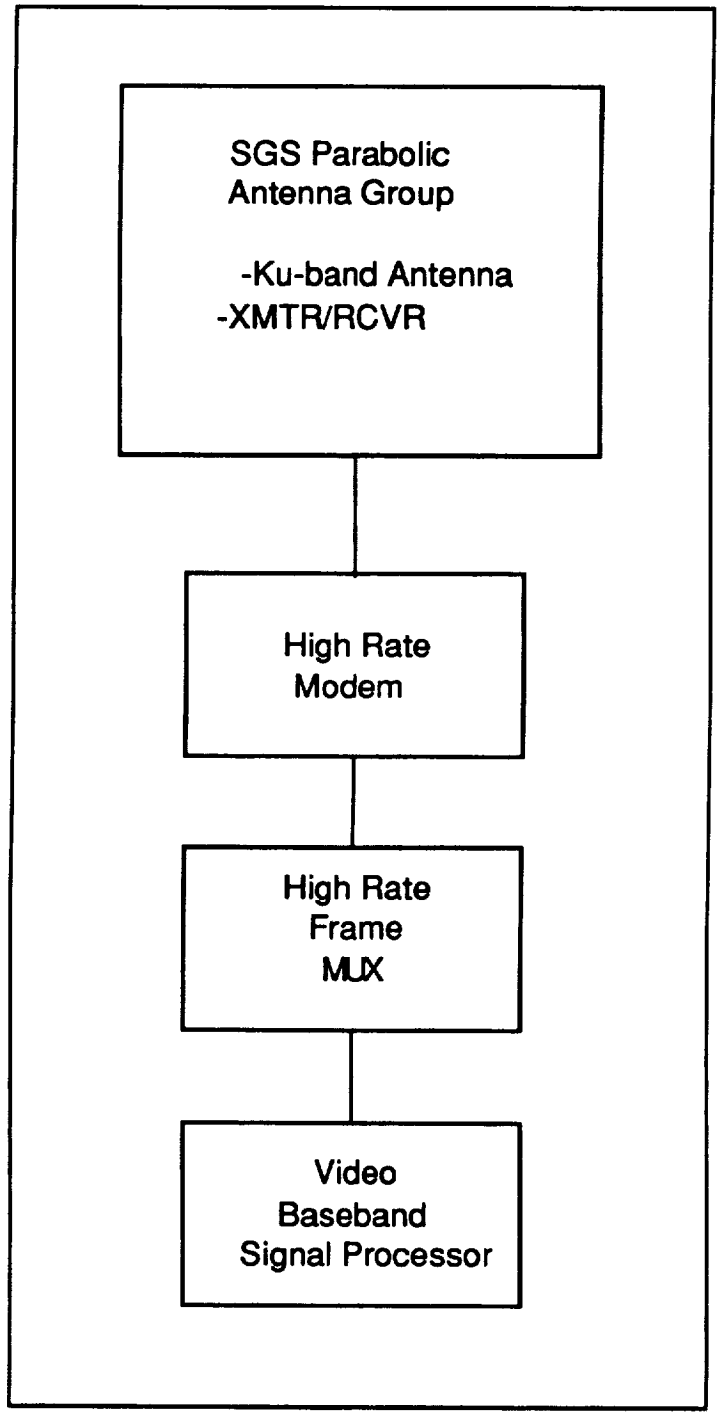


Figure 3.- SGS Hardware Schematic

- Software modeling and evaluation of antenna management (positioning) algorithms.
- Analysis and investigation of selected variables and parameters of these antenna management algorithms, i.e., descriptions and definitions of ranges, scopes, and dimensions.

In a related activity, to assist those responsible for monitoring the development of this flight system software, a brief summary of software metrics concepts, terms, measures, and uses was prepared.

DISCUSSION

ACS Modeling and Analysis

The discussions which follow require the following definitions:
- Rectangular coordinates - coordinates expressed in terms of an x-, y-, and z-component. For example, the TDRSS vectors provided as inputs to the ACS High Gain Antenna Management (ACAM) are given in this coordinate reference. In terms of SSF normal flight mode, defined as "...having the Space Station transverse boom perpendicular to the orbital plane..." [5], the definitions of the x-, y-, and z-axes are as follow:
  - +X-axis - in the direction of travel (and the velocity vector);
  - +Y-axis - out the starboard (right) side of the transverse boom;
  - +Z-axis - towards the center of the earth.
- Spherical coordinates - coordinates expressed in terms of two angles. For the ACS, these two angles are theta and phi, representing the elevation and azimuth position of the TDRS satellite with respect to the SSF, defined as:
  - Elevation - rotation measured with respect to the -Z-axis, with 0 degrees Elevation set at the -Z-axis, and positive elevation measured as motion towards the +X-axis.
  - Azimuth - rotation measured with respect to the +X-axis , with 0 degrees azimuth set at the +X-axis, and positive azimuth measured as motion towards the +Y-axis.
- Unit vector - a vector whose magnitude is equal to unity. For example, the TDRSS pointing vector is a unit vector whose components represent the direction of the TDRSS satellite with respect to the SSF.

The ACAM subcapability includes the following functions related to antenna pointing, an updating activity to be performed once every five seconds to maintain TDRSS contact:
- Accept TDRSS (also referred to by using only the first four initials - TDRS) pointing and rate of change vectors, provided as inputs from the Guidance, Navigation and Control/Propulsion (GN&C/P) System, in rectangular coordinates having x-, y-, and z-components.
- Apply a static bias correction to these TDRSS vectors to compensate for the actual boresight alignment of the S-band horn antennas.

- Convert the resultant vectors from rectangular (x,y,z) coordinates to spherical coordinates (theta, phi) corresponding to the antenna gimbal movements in elevation and azimuth.

- Adjust the resultant antenna gimbal parameters (elevation, elevation rate, azimuth, azimuth rate) so as not to exceed specified software or hardware limits for each.

- Provide the final results for S-band antenna gimbal angles and rates as outputs to be used by the Antenna Controller module to actually drive the gimbal motors to effect motion.

The C&TS FSSR describes algorithms, antenna limits, static bias corrections, and other values required to carry out the ACAM computations. These specifics were used to construct an object-oriented software model using Turbo C++ (Turbo C++ is a registered trademark of Borland International, Inc.). Inputs to the model, representing TDRSS rectangular pointing and rate vectors, were created by one of two methods: (i) using random number generation techniques to derive both unit pointing and constrained rate vectors; or (ii) combining specific unit pointing vector values supplied by a NASA contractor (Lockheed Engineering Services Corporation) with constrained rate vectors derived from modeling work done by a second NASA contractor (Spar Aerospace Limited). [6][7] The Turbo C++ model's intermediate and final results were captured for analysis and verification of inputs and outputs produced by the ACAM process; Table 1 contains an example of model outputs.

The model described above was exercised repeatedly during the research period. A significant number of constants, variables, and parameters specified in the C&TS FSSR were found to have questionable values, ranges, or limits. Consequently, much of the research activity was devoted to investigation of these items, with the following results.

1) The ACS Static Bias Correction - Default Values given in [4], Table 3.2.3-29a, were identified as being 180 degrees out of phase; i.e., the Gamma correction value specified (z-axis rotation) as -PI/4 (-45 degrees) should be given as +3PI/4 (+135 degrees). This error resulted from a choice of coordinate system used in a subcontractor drawing/documentation.

2) The 'Data Type, Units, and Range' characteristics of the 'TDRS n Line-of-Sight' unit vector specified in [3], Table 3.2.3-30, titled 'ACS High Gain Antenna Management Capability Inputs from External systems', were found to be inconsistent. Specifically, the identification of '-1.0E7 to +1.0E7 rad' as the range of values for this unitless pointing vector was both logically and mathematically incorrect; the correct range is given as '-1.0 to +1.0'. And, given this error, the specifications of Accuracy and Precision characteristics are also inconsistently given, both in size and dimensional units.

TABLE 1.- SAMPLE ACS MODEL OUTPUTS

TDRS Raw Data

Position Vector	Rate Vector
X: 0.856803	X: -0.000061
Y: 0.156809	Y: -0.000183
Z: 0.491223	Z: 0.000153

TDRS Biased Data

Position Vector	Rate Vector
X: -0.494970	X: -0.000086
Y: -0.716732	Y: 0.000173
Z: 0.491223	Z: 0.000153

ACS_Gimbal Angle and Rates -- Calculated

Theta (in radians):	2.084290	Elevation (in degrees):	119.420993
Theta-prime (in radians):	0.000175	Elev'n rate (in degrees):	0.010038
Phi (in radians):	-2.175181	Azimuth (in degrees):	-124.628716
Phi-prime (in radians)	-0.000194	Azimuth rate (in degrees):	-0.011126

ACS_Gimbal Angle and Rates -- Clipped

Theta (in radians):	2.084290	Elevation (in degrees):	119.420993
Theta-prime (in radians):	0.000175	Elev'n rate (in degrees):	0.010038
Phi (in radians):	-2.175181	Azimuth (in degrees):	-124.628716
Phi-prime (in radians)	-0.000194	Azimuth rate (in degrees):	-0.011126

3) The 'Data Type, Units, and Range' characteristics of the 'TDRS n Line-of-Sight Rate' vectors specified in Table 3.2.3-30 of [3] were determined to be in error. The components of these vectors represent instantaneous velocities along the three axes and thus require that negative magnitudes be included.

The associated TDRS Line-of-Sight unit vectors contain dimensionless (unitless) values representing the relative position, not range, of the TDRS and the SSF. Thus, use of 'rad/sec' with respect to changes in these unitless vectors, given in rectangular coordinates, is logically inconsistent. Moreover, this inconsistency can be confirmed by examining the TDRSS rate vector mathematical formulations found in [8]. The dimensional units for this input parameter are correctly given as sec^{-1}.

The Range magnitude given as 0..0.1 (rad/sec) is also believed to be incorrect; the origin of this range value could not be identified. However, review of SSF subcontractor system documentation and discussion with several

subcontractor representatives suggests a more representative range for the components of this rate vector would be -2.0E-4 to +2.0E-4 sec^{-1}, calculated given SSF and TDRSS nominal orbits and attitudes.

4) There appears to be a general inconsistency, or at least an opportunity for confusion, with respect to the number of TDRSS satellites. In [3], Table 3.2.3-30, for example, one finds reference to 'TDRS One ' through 'TDRS Six'. In Table 3.2.3-25, one finds reference to 'TDRS Occultation 1' through 'TDRS Occultation 4'. And in other parts of the C&TS FSSR, there are references to TDRS-East and TDRS-West. It would be useful to reconcile these differences for the reader of the FSSR.

5) The values specified as 'Azimuth' and 'Elevation' in Table 3.2.3-30, inputs which define the ACS antenna 'park' position, appear to be consistent with the electrical and software limits given earlier in the ACAM section. However, comparing these values to the 'ACRFG TDRS Pointing Vectors Control' values given in Table 3.2.3-68-7C.1, there are (unresolved) inconsistencies for these parameters.

6) Table 3.2.3-68-7.C1 also gives ranges for 'TDRS Azimuth Angular Velocity' and 'TDRS Elevation Angular Velocity' which appear to be in error. In [7], velocity constraints are given as 9 degrees per second for azimuth and elevation, values significantly larger than the -0.10 to +0.10 degree/sec specified in this table. Similarly, then, there is reason to question the associated Precision attributes shown for the angular velocity entries of Table 3.2.3-68-7C.1.

SGS Modeling and Analysis

The discussions which follow require the following definitions:

- Elevation (pitch) - rotation measured with respect to the -Z axis (about the y-axis), with 0 degrees elevation set at the -Z axis, and positive elevation measured as motion towards the +X axis.
- Cross-elevation (roll; or azimuth) - rotation measured with respect to the -Z axis, with 0 degrees azimuth set at the +X axis, and positive azimuth measured as motion towards the +Y axis.

The SGS Antenna Management (SGAM) capability includes the following functions related to antenna pointing, an updating activity to be performed at least once per second.

- Accept TDRSS pointing and rate of change vectors, provided as inputs from the GN&C/P system, in rectangular coordinates.
- Convert the TDRSS vectors from rectangular to spherical coordinates (alpha, beta) corresponding to the SGS antenna parameters of pitch and roll (or elevation and cross-elevation; or elevation and azimuth). This translation between coordinate systems directly incorporates static bias correction to compensate for antenna alignment.
- Adjust the resultant antenna parameters (elevation, elevation rate, azimuth, azimuth rate) so as not to exceed specified software or hardware limits.

- Provide the final results for KU-band antenna angles and rates as outputs to be used for actual antenna positioning.

The C&TS FSSR describes algorithms, antenna limits, and other specifics required to carry out SGAM computations. The Turbo C++ model used for the ACS (ACAM) evaluation, described earlier, was modified and adapted to develop an equivalent SGS (SGAM) model. Input were identical to those used for ACS modeling; an example of model outputs is shown in Table 2.

TABLE 2.- SAMPLE SGS MODEL OUTPUTS

TDRS Raw Data

Position Vector	Rate Vector
X: 0.856803	X: -0.000061
Y: 0.156809	Y: -0.000183
Z: 0.491223	Z: 0.000153

SGS_Gimbal Angle and Rates -- Calculated

Alpha (in radians):	-1.050225	Elevation (in degrees):	-60.173452
Alpha-prime (in radians):	0.000163	Elev'n rate (in degrees):	0.009324
Beta (in radians):	0.157459	Azimuth (in degrees):	9.021721
Beta-prime (in radians)	-0.000036	Azimuth rate (in degrees):	-0.002088

SGS_Gimbal Angle and Rates -- Clipped

Alpha (in radians):	-1.050225	Elevation (in degrees):	-60.173452
Alpha-prime (in radians):	0.000163	Elev'n rate (in degrees):	0.009324
Beta (in radians):	0.157459	Azimuth (in degrees):	9.021721
Beta-prime (in radians)	-0.000036	Azimuth rate (in degrees):	-0.002088

The SGS model was exercised a limited number of times due to the research emphasis placed on the more critical ACS system. Nonetheless, this limited SGS modeling and analysis highlighted inconsistencies similar to those found for the ACS. Specifically, the following areas were identified.

1) Specification and definition of 'TDRS n Line of Sight' and 'TDRS n Line of sight Rate' vectors found in [4], Table 3.2.4-38 titled 'SGS Antenna Management Capability Inputs from External Systems', included inconsistencies similar to those discovered for the ACS's ACAM subcapability.

X-, y-, and z-components of the TDRS Line of Sight vectors are erroneously characterized as meters, with an erroneous range of -1E7 to +1E7 meters. These components are unitless and lie between -1.0 and +1.0.

X-, y-, and z-components of the TDRS LIne of Sight Rate vectors are erroneously specified in terms of radians per second. The 'Range' value given, -0.1 .. +0.01 radians per second, obviously has the wrong dimensions; the correct magnitude of this range is unknown.

The 'TDRS Line of Sight Rate Vector' issues discussed in the ACS Modeling and Analysis section apply equally to the SGS system.

2) The specification and description of 'SGS Antenna Management Capability Outputs to External Systems', given in Table 3.2.4-40 of [4], merit further analysis and exploration. It appears there are inconsistencies in this table of SGS parameters similar in number and scope to those found in the corresponding ACS (ACAM) table.

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

Software modeling and evaluation of ACS and SGS antenna management (positioning) algorithms have revealed numerous discrepancies in the documentation and description of their capabilities and requirements.

A significant number of the characteristics of the ACS and SGS subsystems, i.e., constants, variables, and parameters, appear to be logically or mathematically inconsistent. The utility of the software modeling research activity itself derived from the identification of such uncertainties, and the opportunity to resolve some of them.

There still exists almost universal misunderstanding, or at least disagreement, surrounding the definition and description of the TDRSS Line-of-Sight Rate Vectors, so much so in fact, that these rate vectors, their interpretation, composition, and use deserve continued research and evaluation. These rate vectors play significant roles in the actual gimbal motor drive algorithms [9], so they need to be well understood, a condition that certainly is not true at this moment.

Recommendations

Complete and detailed mathematical analysis of the ACS, then the SGS, antenna management (positioning) algorithms should be completed to establish a solid foundation for the description of these subsystems and their capabilities in the C&TS FSSR.

Software modeling of the ACS and SGS subsystems should continue and should incorporate orbital equations and other mathematical information from which TDRSS Line-of-Sight and TDRSS Line-of-Sight Rate vectors can be numerically produced, thus replicated. Software modeling should continue to be viewed as a vehicle for asking -- and answering -- questions related to the ACS and SGS subsystem components and functional requirements.

Contract should require a common coordinate reference system be used by all participants for all documents and contract submissions.

Contracts should require documentation, derivation, and/or direct references for each parameter (e.g., range; dimensional units; magnitudes) used in the description and characterization of system capabilities such as ACAM and SGAM.

REFERENCES

[1] *Block Update to C & T Architectural Control Document, SSP 30260.* Space Station Control Board Directive No. BB000593, 10/24/89.

[2] *Contract End Item (CEI) Specification for Communication and Tracking System (C & T) (CEI Number 212002A), Work Package 2 (WP-2).* Contract NAS 9-18200, Specification Number SP-M-002E, dated 27 May 1992.

[3] *Flight System Software Requirements (COMMUNICATIONS AND TRACKING SYSTEM SOFTWARE- Part 2, Section 3.2 through 3.3 (Level C), SSP 30606.* Contract NAS 9-182000. McDonnell-Douglas Aerospace Division. Volume II, Revision B, dated 15 June 1993 (DRAFT REVIEW COPY). Section 3.2.3.3.2.1.4: ACS Service Maintenance (ACSM), pp 3.2.3-87 through 3.2.3-96; and Section 3.2.3.3.2.1.6: ACS High Gain Antenna Management (ACAM), pp 3.2.3-98 through 3.2.3-110, and page 3.2.3-247.

[4] *Flight System Software Requirements (COMMUNICATIONS AND TRACKING SYSTEM SOFTWARE- Part 2, Section 3.2 through 3.3 (Level C), SSP 30606.* Contract NAS 9-182000. McDonnell-Douglas Aerospace Division. Volume II, Revision B, dated April 5, 1993. Section 3.2.4.3.2.1.6: SGS Antenna Management (SGAM), pp 3.2.4-182 through 3.2.4-217.

[5] *SSP Assembly Contingency System Control Algorithms: Final Design Review Package, RML-009-93-034.* King, D., Spar Aerospace Limited, Quebec, November, 1992.

[6] *Space Station and TDRSS Orbital Model, Project No: 5209-A, Report No: RML-009-92-152.* Arbery, G., Spar Aerospace Limited, Quebec, November, 1992.

[7] *ACS RF Group CDR Technical Summary, Spar Program No. 5209-A, RML-009-93-093.* Gratton, D., Spar Aerospace Limited, Quebec, undated.

[8] *Flight System Software Requirements (GUIDANCE, NAVIGATION & CONTROL/PROPULSION), Part 9, Appendix I.40, SSP 30606.* Contract NAS 9-182000. McDonnell-Douglas Aerospace Division. Volume I, Revision A, dated April, 1993. Section I.40.2.3: TDRS Line-of-Sight and Line-of-Sight Rate, pp I.40.2-17 through I.40.2-19.

[9] *ACS ACON Software Design Document (SDD), Spar Aerospace Limited Project No. 5209A, Drawing No. 827009,* pp 88-95. Spar Aerospace Limited, Quebec, December 7, 1992.

N94-25375

INVESTIGATING PYROLYSIS/INCINERATION AS A METHOD OF RESOURCE RECOVERY FROM SOLID WASTE

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Bobby J. Roberson, Ph.D., Christopher S. LeMay
Academic Rank:	Professor, Student
College and Department:	Bevill State Community College Chemistry Department Fayette, AL 35555
NASA/JSC	
Directorate:	Engineering
Division:	Crew and Thermal Systems
Branch:	Life Support Systems
JSC Colleague:	Charles E. Verostko
Date Submitted:	August 6, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

Pyrolysis/incineration (P/I) is a physicochemical method for the generation of recoverable resources from solid waste materials such as inedible plant biomass (IPB), paper, plastics, cardboard, etc. P/I permits the collection of numerous gases with a minimal amount of solid residue. Pyrolysis, also known as starved air incineration, is usually conducted at relatively high temperatures (> 500 degrees Celsius) in the absence of oxygen. Incineration is conducted at lower temperatures in the presence of oxygen. The primary purpose of this study was to design, construct, and test a model P/I. The system design includes safety requirements for temperature and pressure. The objectives of this study were: 1) to design and construct a P/I system for incorporation with the Hybrid Regenerative Water Recovery System ; 2) to initiate testing of the P/I system; 3) to collect and analyze P/I system data; 4) to consider test variables; and 5) to determine the feasibility of P/I as an effective method of resource recovery. A P/I system for the recovery of reuseable resources from solid waste materials was designed, constructed, and tested. Since a large amount of inedible plant biomass (IPB) will be generated in a space-based habitat on the lunar surface and Mars, IPB was the primary waste material tested in the system. Analysis of the effluent gases was performed to determine which gases could be used in a life support system.

INTRODUCTION

Several waste recovery technologies are being considered for use on lunar and Mars habitats. Some of these are biological while others are physicochemical. Others are even hybrids of the preceding types. Results of a NASA Workshop indicate that several waste recovery technologies must be considered and pursued for the development of a human-rated test facility (5). Two systems will then be selected for lunar and Mars Controlled Ecological Life Support Systems (CELSS) applications. Design, construction, and testing of the two prototypes and production of two flight systems should be complete and ready by a target flight date early in the twenty-first century. One of the physicochemical technologies that offers promise for suitable use is pyrolysis/incineration (P/I). Literature reviews and experimentation indicate that a P/I system can be safe, efficient, and practical and might also have Earth-based applications for the conversion of solid waste materials into reuseable resources (4).

One of the major objectives of this study was to compare the processes of combustion and P/I. Recent investigations involving the combustion of IPB reveal that the conversion of IPB to carbon dioxide and water requires about the same amount of oxygen that was produced by photosynthesis. In other words, the amount of oxygen produced during crop growth is equal to the amount required for the oxidation of the biomass produced during plant growth; therefore, the actual amount of oxygen produced that is available for human consumption is in proportion to the amount of biomass actually utilized by humans. The remaining oxygen must be available to oxidize the rest of the biomass back to carbon dioxide and water or the system will not be a regenerative one (2).

Since the IPB is composed of those plant parts that are not used for human consumption such as roots, stems, leaves, etc., these components represent a significant portion of the total plant mass. Specifically, the IPB for lettuce is approximately 10%, 65% for soybeans, and 60% for wheat (2). It is also important to note that not all of the edible plant biomass (EPB) is digestable and will require additional oxidation outside of the humans before it can be recycled back to plant production. A waste model has been developed based upon information from a NASA Workshop (5). Workshop data indicate that the amount of waste generated per person per day in a space-based habitat will be: 1.85 kg of IPB, 0.9264 kg of trash, 0.194 kg of paper, and 0.041 kg of filters. In a space-based CELSS, it is

desirable that the maximum amount of biomass produced be utilized by the crew so that mass, volume, and energy requirements be minimized. Any biomass oxidation process outside of the crew food cycle reduces the overall efficiency of the CELSS in terms of direct crew life support.

Incineration

Incineration is the oxidation of waste using either limited amounts of pure oxygen or oxygen diluted with an inert gas, usually argon or nitrogen, whereas combustion is the complete oxidation of a substance in an excess of oxygen or air. A temperature of 1000 degrees Celsius and a pressure of one atmosphere are preferred. The amount of air present is in 75 to 100% excess (3). The process is relatively fast and results in 97 to 99% reduction of waste materials such as IB, paper, plastics, food and human waste, etc. the excess air, turbulence, and retention time provide conditions for complete combustion. The chemical reaction for the incineration of cellulosic materials is:

$$C_6H_{12}O_5 + 6\,O_2 + \text{heat} \quad \text{----->} \quad 6\,CO_2 + 5\,H_2O$$

Pyrolysis

Pyrolysis is the chemical destruction of a carbonaceous material by heating to 400 to 1000 degrees Celsius in the absence of oxygen, or in a controlled oxygen environment. The process usually results in the formation of four phases of products: a) a gaseous component including hydrogen, methane, carbon monoxide, and carbon dioxide; b) an oil including organic acids, alcohols, ketones, etc.; c) a char including carbon and inert materials; and d) an aqueous phase containing some water-soluble compounds. In general, rapid heating results in higher yields of gases, whereas slower heating results in higher amounts of oils and char (1). The overall reaction for the pyrolysis of cellulosic materials is:

$$C_6H_{12}O_5 + \text{heat} \quad \text{----->} \quad CH_4 + 2\,CO + 3\,H_2O + 3\,C$$

Depending on the temperature of the reaction, the pressure in the chamber, and the amount of air (or oxygen) permitted to enter the chamber, the reaction products will vary in composition.

DISCUSSION

Pyrolysis of IPB and paper revealed that several reuseable gases are produced. The process also reduced the mass of solid waste by 97 to 99%. The P/I process was performed in an open system at ambient pressure. It is believed that a combination of high temperature and high pressure might result in a higher degree of decomposition of the waste.

Figure 1 is a schematic of the P/I process.

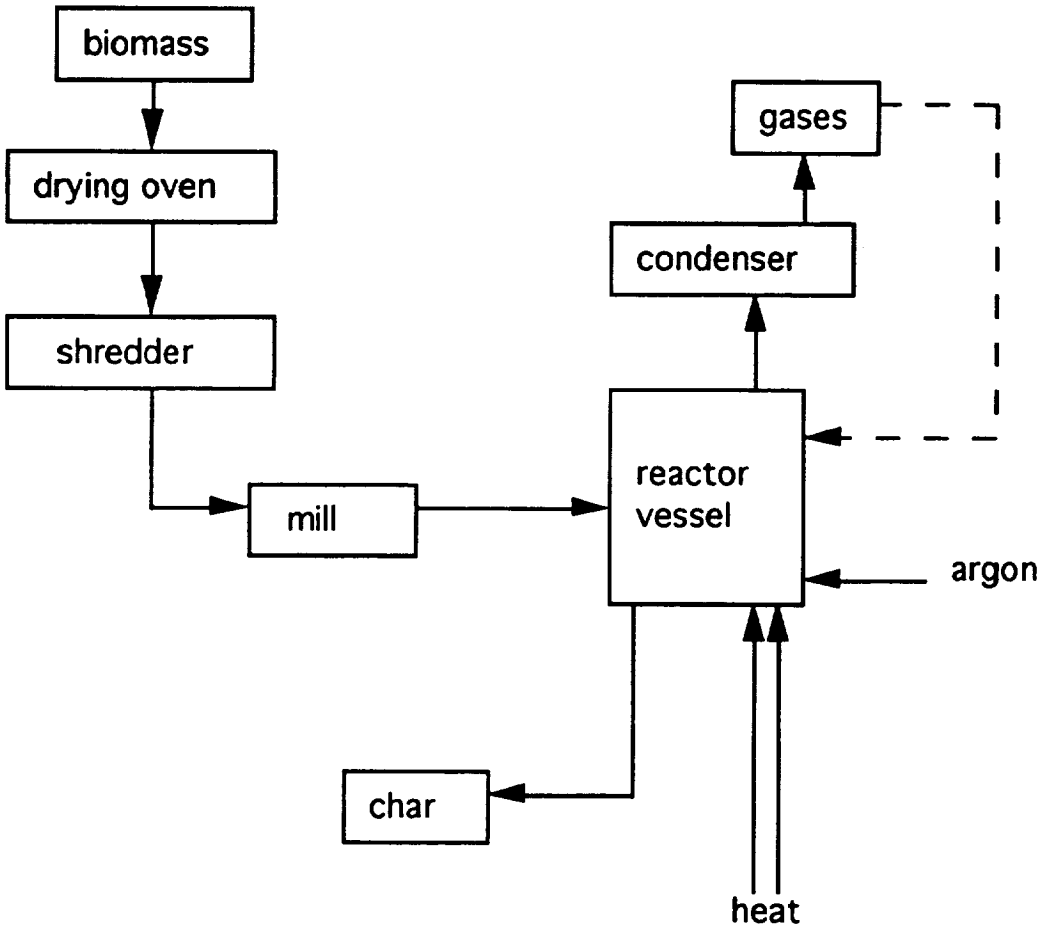


Figure 1.- Schematic of Pyrolysis/Incineration System

A diagram of the pyrolysis reactor system is shown in Figure 2. (This system was constructed by Benjamin P. Fowler, Fowler Engineering Company, Houston, TX.)

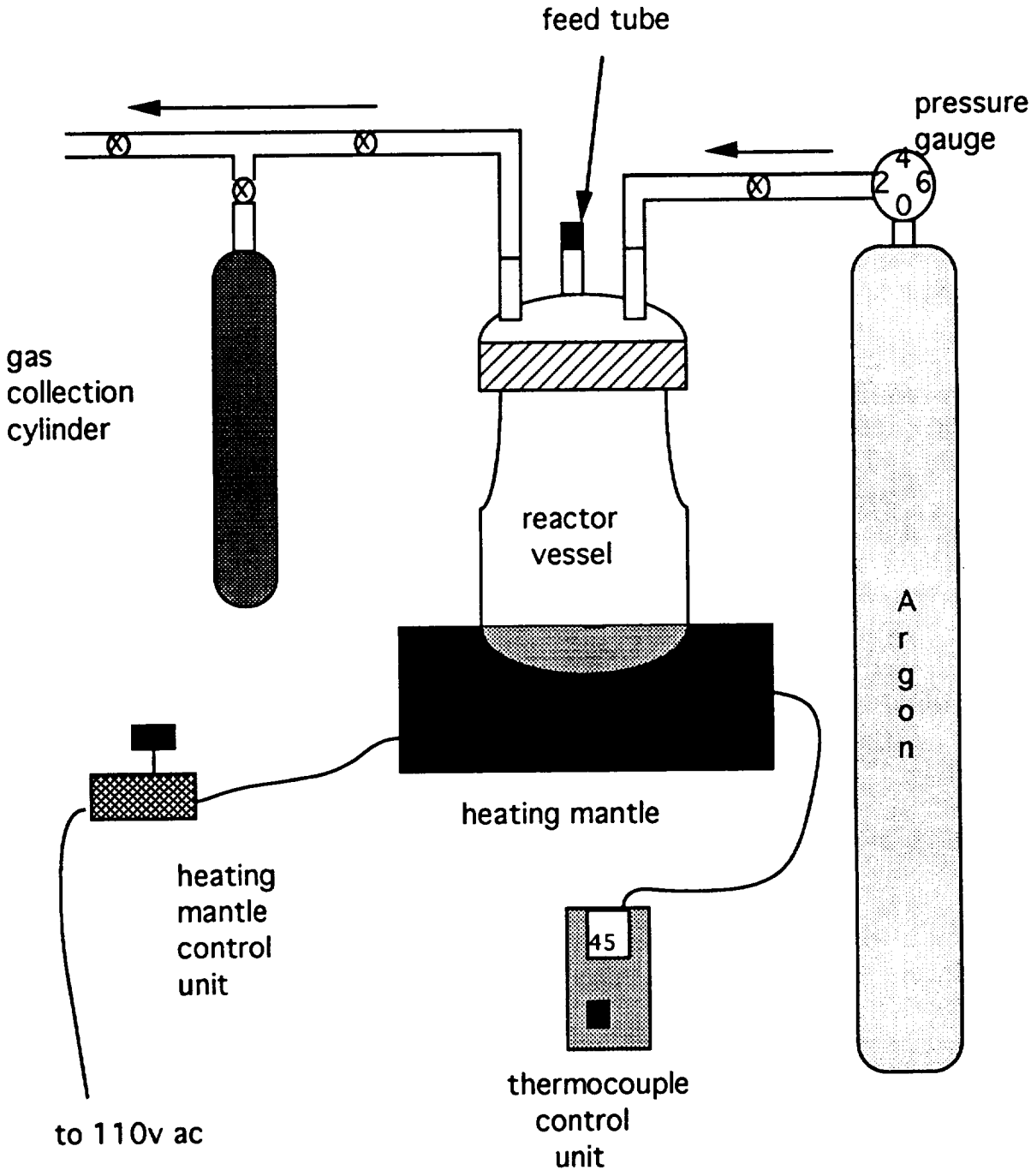


Figure 2.- Diagram of the Pyrolysis System

System Description

The reactor vessel had a volume of 1.35 liters. The base of the reactor vessel was constructed of #314 stainless steel. The base consisted of a 4" weld cap welded to a 4" by 3" reduction section. The reduction section was welded to a 3" by 3" nipple. The head was constructed of #413 stainless steel. It consisted of a 3" weldcap with a 3" threaded coupling. The inlet/ outlet tubes were 1/4" by 4" stainless steel threaded tubes. A GLAS-COL Model # STM 900 heating mantle, a CORD TROL Model # PL-112 heater control unit, and an Atkins K Model # 39658-K thermocouple thermometer were used.

Procedures

Sample Preparation. Samples of IPB (wheat straw and roots) were dried in an oven at 75 degrees Celsius for 48 hours. Samples were shredded and then passed through a 40-mesh screen using a Wiley mill. Samples of paper were cut into 3 mm squares.

Gas Collection. Valves on evacuated, stainless steel, gas collection tubes were opened at specified temperatures and were left open for approximately 5 minutes. After 5 minutes, the valves were closed, and the gas collection tubes were replaced.

Gas Analysis. Gases were analyzed using a Hewlett Packard 5880 Gas Chromatograph and a Hewlett Packard 5987 Gas Chromatograph/Mass Spectrometer.

Residue Collection. After cooling, the reactor vessel was opened. The residue was collected and weighed.

Experimental Results

The experimental results obtained from the pyrolysis of various waste materials are shown in Tables 2 and 3. Table 2 indicates that pyrolysis is very effective in reducing the mass of cellulosic waste material. Table 3 reveals that several short-chain hydrocarbons are porduced when cellulosic materials are pyrolyzed. These hydrocarbons have some value as fuels. It is believed that the amount of fuel-grade products might be increased by using a high pressure and high temperature system and by using samples that have large surface areas.

TABLE 1.- SOLID MASS REDUCTION BY PYROLYSIS*

	SAMPLES	
	wheat, 40-mesh	paper
initial mass (g)	1.005	1.110
final mass (g)	0.073	0.052
% Reduction	92.8	95.3

* Pyrolysis conditions were 767 degrees Celsius for 30 minutes.

TABLE 2.-MAJOR PYROLYSIS PRODUCTS[1] (in ppm)

	Samples		
Gases	Wheat[2]	Wheat[3]	Paper[3]
methane	35.8	40.0	116.0
ethylene	9.0	11.0	23.0
ethane	4.1	6.0	8.0
propylene	8.0	4.0	16.0
propane	33.2	7.0	2.0

[1] Trace components included primarily acetaldehyde and C-3 through C-5 aldehydes and ketones.
[2] A slow heating process was used (ambient to 767 degrees Celsius in 30 minutes).
[3] A rapid heating process was used (ambient to 767 degrees Celsius in 15 minutes).

Specifications for Flight-Ready System

The two most important factors to be considered for a flight-ready system are the use of high temperature and high pressure. A system used in a space-based habitat should also be relatively small and have low power requirements. The system should also be easily maintained. With these objectives and factors in mind, the specifications shown in Table 3 are recommended for a flight-ready system. A suggested design for the reactor vessel is shown in Figure 3. (This system is similar to the Parr high pressure bomb Model 4680 described in the Parr Instrument Catalog, 7th Ed., 1991, pp. 54-55 and 106. The esitmated cost of the system is

approximately $15,000.) The proposed system will be capable of combustion, incineration, and pyrolysis processes.

TABLE 3.- SPECIFICATIONS FOR A FLIGHT-READY PYROLYSIS/INCINERATION SYSTEM

Volume	2000 ml
Sample mass	500 g
Maximum operating temperature	600 degrees Celsius
Maximum operating pressure	4000 psi
Construction material	stainless steel
Feed mode	batch
Heater	Electric element embedded in a shaped ceramic body
Head openings	gas inlet, gas outlet, feed inlet

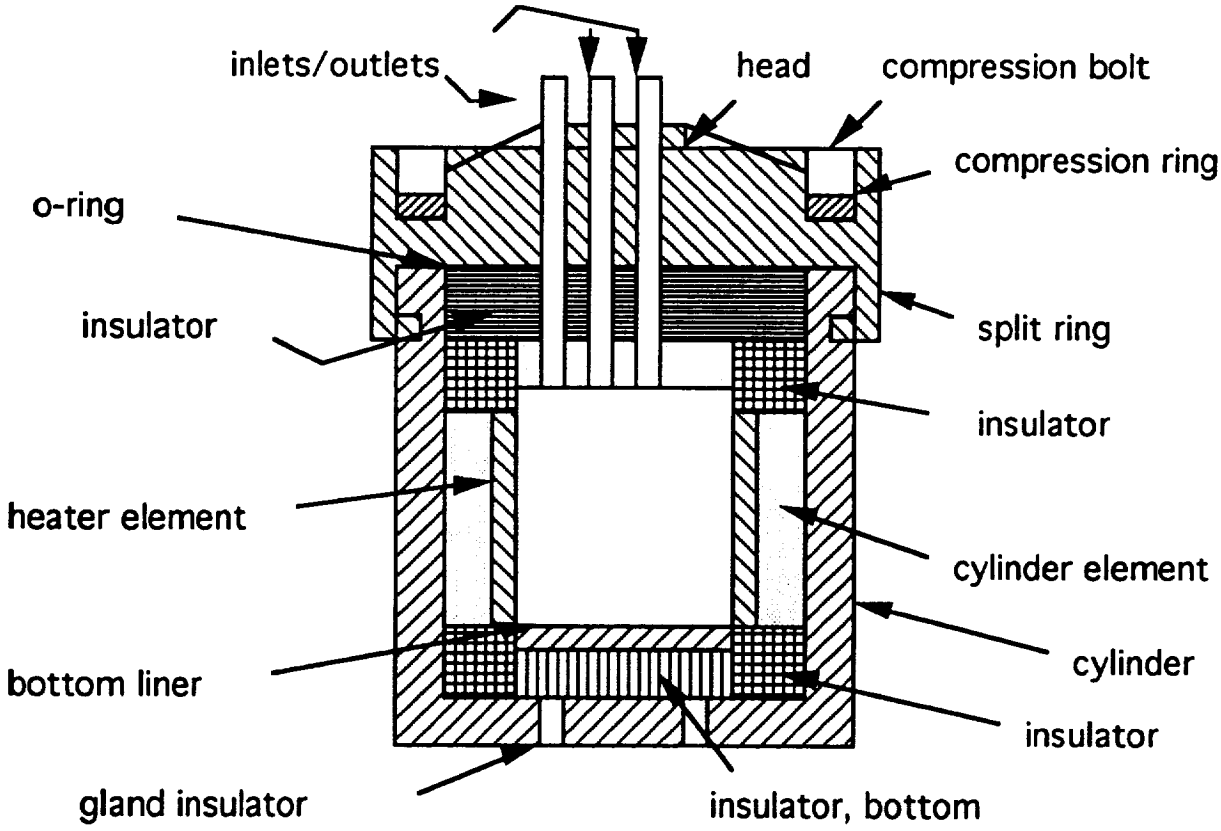


Figure 3.- Diagram of Proposed Pyrolysis Reactor

CONCLUSIONS

The results of this study indicate that pyrolysis/incineration is a physicochemical technology that has practical applications for space-based habitats as well as for waste management and resource recovery on the Earth. An appreciable reduction in the mass of the solid waste plus the recovery of reuseable gases from a variety of waste materials was achieved in a pyrolysis/incineration process. Due to the limited quantities of fuel-grade gases produced, it is believed that these gases should only be used to heat the pyrolysis reactor vessel. The effluent gases contain some particulate matter and numerous trace contaminants such as organic acids, ketones, and aldehydes. These substances should probably be removed if the fuel-grade gases were to be used at remote locations. Any carbon dioxide produced could be used in a plant growth chamber. The major problem with a pyrolysis process is the limited amount of sample that can be treated at high temperature in a closed vessel due to the high pressure generated. Possibly an open system with a continuous injection feed mechanism could offer a solution to this problem.

REFERENCES

1. *Combustion and Incineration Processes.* W. R. Niessen, Ed. 1978. Marcel Dekker.

2. Dreschel, T. W., R. M. Wheeler, C. R. Hinkle, J. C. Sager, and W. M. Knott. "Investigating Combustion as a Method of Processing Inedible Biomass Produced in NASA's Biomass Production Chamber. NASA Tech. Memo. 103821. May, 1991.

3. *Incineration Processes.* C. R. Brunner, Ed. 1984. Van Nostrand-Reinhold Co.

4. Roberson, B. J. "Development of a Pyrolysis Waste Recovery Model with Designs, Test Plans, and Applications for Space-Based Habitats." Final Report. NASA/ASEE Summer Faculty Fellowship Program. August, 1992.

5. Verostko, C. E., N. J. C. Packham, and D. L. Henninger. Final Report on NASA Workshop on Resource Recovery from Wastes Generated in Lunar/Mars Controlled Ecological Life Support Systems (CELSS). JSC Doc. No. 25736, CTSD-ADV-035, August, 1992.

N94-25376

DEVELOPMENT OF A MODEL TO ASSESS ORTHOSTATIC RESPONSES

Final Report

NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared by:	Marilyn Rubin, Ph.D.
Academic Rank:	Professor
University & Department:	Saint Louis University School of Nursing St. Louis, Missouri 63104
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Biomedical Laboratories
JSC Colleague:	Suzanne M. Fortney, Ph.D.
Date Submitted:	October 8, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

A major change for crewmembers during weightlessness in microgravity is the redistribution of body fluids from the legs into the abdomen, thorax, and head. The fluids continue to be sequestered in these areas throughout the flight. Upon reentry into gravity on landing, these same body fluids are displaced again to their normal locations, however, not without hazardous incidence to the crewmembers. The problem remains that upon landing, crewmembers are subject to orthostasis, that is, the blood flowing into the legs reduces the blood supply to the brain and may result in the crewmember fainting.

The purpose of this study was to develop a model of testing orthostatic responses of blood pressure regulating mechanisms of the cardiovascular system, when challenged, to maintain blood pressure to the brain. To accomplish this, subjects responses were assessed as they proceeded from the supine position to progressive head-up tilt positions of 30°, 60°, and 90° angles.

A convenience sample consisted of 21 subjects, females (N=11) and males (N=10), selected from a list of potential subjects available through the NASA subject screening office. The methodology included all non-invasive measurements of blood pressure, heart rate, echocardiograms, cardiac output, cardiac stroke volume, fluid shifts in the thorax, ventricular ejection and velocity times, and skin blood perfusion.

The Fischer statistical analysis was done of all data with the significance level at .05. Significant differences were demonstrated in many instances of change of posture for all variables. Based on the significance of the findings of this study, this model for assessing orthostatic responses does provide an adequate challenge to the blood pressure regulatory systems. While individuals may use different adaptations to incremental changes in gravity, the subjects, in aggregate, demonstrated significant adaptive cardiovascular changes to orthostatic challenges which were presented to them.

INTRODUCTION

A major change for crewmembers during weightlessness in microgravity is the redistribution of body fluids from the legs into the abdomen, thorax, and head. The fluids continue to be sequestered in these areas throughout the flight. Upon reentry into gravity on landing, these same body fluids are displaced again to their normal locations, however, not without hazardous incidence to the crewmembers. Bungo (1989) observed that upon assuming an upright position on landing, orthostatic intolerance has been consistently observed after space flight. Greenleaf et al (1989) have recognized that physical exercise, pre-reentry fluid loading of the crew, and G-suit inflation have been used as countermeasures to maintain orthostatic tolerance. Melchior and Fortney (1993) described another countermeasure of lower body negative pressure (LBNP) applied during flight as another methodology to improve orthostatic tolerance of the crew upon landing of the orbiter. While there has been a modicum of success with these various protocols, none has successfully solved the problem of orthostatic intolerance of the crew. Since orthostasis can result in fainting, the safety of the crew and, especially, their ability to egress the orbiter in an emergency, becomes a priority. Refinement of these countermeasures previously described continues with ground based bedrest studies and in flight weightlessness studies.

The purpose of this study was to develop a model of testing orthostatic responses of the blood pressure regulating mechanisms of the cardiovascular system, when challenged, to maintain blood pressure to the brain. To accomplish this, subjects' responses were assessed as they proceeded from the supine position to progressive head-up tilt positions of 30°, 60°, and 90° angles.

METHODOLOGY

The convenience sample consisted of 21 subjects, females (N=11) and males (N=10), selected from a list of potential subjects which was available at the NASA subject screening office. All subjects were required to have a current Air Force Class III physical examination and screening for drugs and HIV. The subjects were within the age range of 23-51 years (mean= 34 years), a range in height of 61-75 inches (mean=67 inches), and a range of weight of 107-211 pounds (mean= 147 pounds). The subjects were asked to abstain from alcohol ingestion for 48 hours prior to the tests and abstain from caffeine ingestion for 12 hours prior to the test. They were non-smokers and were not taking any prescription drugs.

Upon arrival at the laboratory, the subject changed into shorts, a t-shirt and athletic shoes. Weight and height were measured and recorded. The investigator then requested the subject to lie in a supine position on a circoelectric bed for a period of 30 minutes to establish baseline data at 0°. At 25 minutes of the rest period, data was collected for 5 minutes. At the end of this period, the bed was tilted to a 30° angle for 5 minutes, and successively to 60° and 90° angle (standing position) tilts, each for 5 minutes. The angles correspond to gravity increments (obtained from the sines of the angles) from +0.50 g (30°), +0.87 g (60°), to +1 g (90°). The time to change from one position to another was an average of 6 seconds. The subject was then returned to

the supine position at 0° for five minutes for recovery. Data was collected continuously which included blood pressure (manually), heart rate, electrocardiogram, echocardiogram, fluid shifts within the thorax, and laser doppler measurement of skin blood perfusion of the left forearm and calf of the left leg. All measurements were non-invasive. Each subject spent approximately one hour in testing. The ambient temperature of the laboratory was an average 22° C. This investigator explained the protocol to each subject and informed consent was obtained before the subject's participation.

Blood perfusion and blood flow velocity of the skin of the inner aspect of the left forearm and outer aspect of the left calf of the leg were measured by placing a surface laser probe in both locations. To avoid vasoconstriction from a cool environment, the probes were placed in circular heaters which surrounded the probe with an opening at the bottom for access of the laser to the skin. The heaters were fixed to the skin with adhesive disks. The temperature of the heaters was 38° C. Data was collected continuously on a Perimed Laser Doppler Instrument and recorded simultaneously on a computer with a visual graphic display of the data.

Fischer et al (1986) described laser doppler flowmetry as a method of measurement of red blood cell flux within the capillary bed. As light reflects a moving object, it changes its frequency in proportion to the velocity of the moving object. The light is transmitted to the skin surface through a fiberoptic cable and reflected back through the cable to the instrument for analysis.

Cardiac stroke volume was measured by continuous wave Doppler method. Data was collected through skin electrodes placed on the subject's thorax. The pulse velocity was measured at the suprasternal notch where aortic diameter was measured by m-mode echocardiography. Cardiac output was calculated by multiplying the cardiac stroke volume by the heart rate.

The Biomed Instrument was used to measure thoracic impedance or fluid shifts in the thorax. Two skin electrodes were placed on the subject on either side of the neck and two on the lower chest.

Blood pressure measurements were made manually each minute during the baseline, 5 minute intervals, and at the end of recovery. An aneroid sphygmomanometer and stethoscope were used for collecting this data.

A Finapres instrument was attached to the subject's left middle finger with an inflated cuff. This instrument was used to observe graphically displayed blood pressure trends and digital readout of blood pressure, heart rate, and pulse pressure.

RESULTS

Statistical analyses were done using the Fisher test for significant difference at the .05 level for all variables. For this report, data are presented for 13 subjects because the data analyses is not yet completed for the remaining 8 subjects.

HEART RATE

The mean heart rate ranged from 59 to 75 beats per minute; the stroke volume ranged from 51 to 88 beats per minute; and the cardiac output ranged from 3757 to 5360 ml/minute over the entire testing period.

Mean heart rate was at its minimum at the end of the 5 minute recovery period when the subject was returned to the supine position, which was slightly lower than the mean of the baseline rest period at the beginning of the test. The heart rate increased progressively to its maximum at 90° head-up tilt.

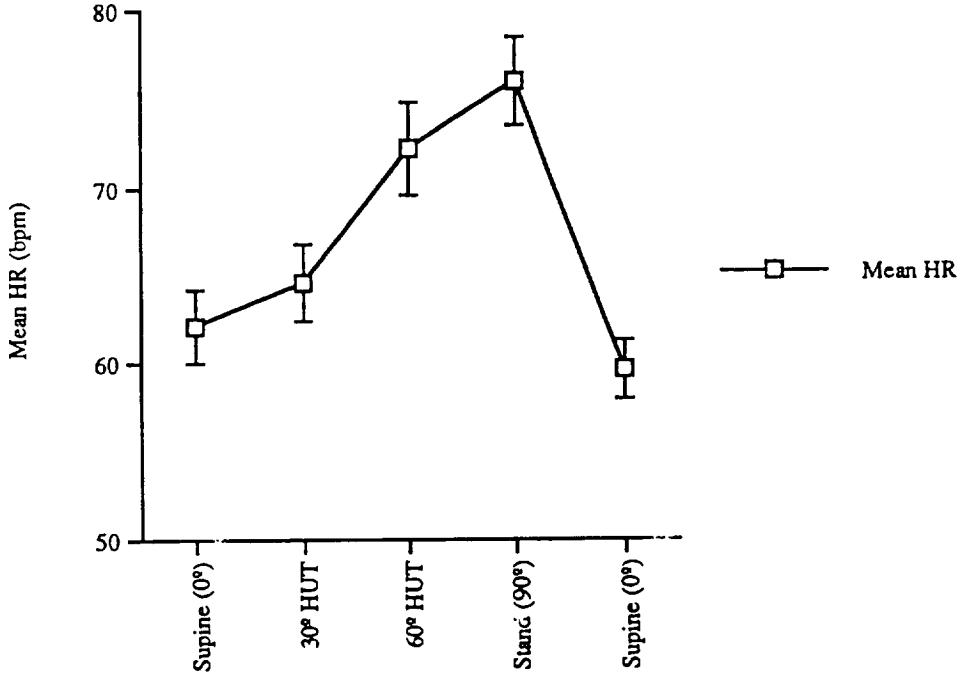


Figure 1. Mean Heart Rate (HR) Responses For All Positions

STROKE VOLUME

The stroke volume reached its maximum level at the end of recovery when the subject was returned to the supine position. This was slightly increased over the baseline mean. The stroke volume progresively decreased at 30°, 60°, and reached a minimum at 90° headup tilt.

CARDIAC OUTPUT

The mean cardiac output reached its highest level during recovery with the subject in supine position at the end of the test. It was slightly higher than at baseline. The cardiac output decreased progressively from the baseline to the 90° headup tilt position when it reached its lowest value.

The cardiac output was significantly different from the supine baseline position to 30°, 60° , and 90° head-up tilt positions; from 30° head-up tilt position to 90° head-up tilt and supine position of recovery; from 60° and 90° head-up tilt positions to the supine position of recovery.

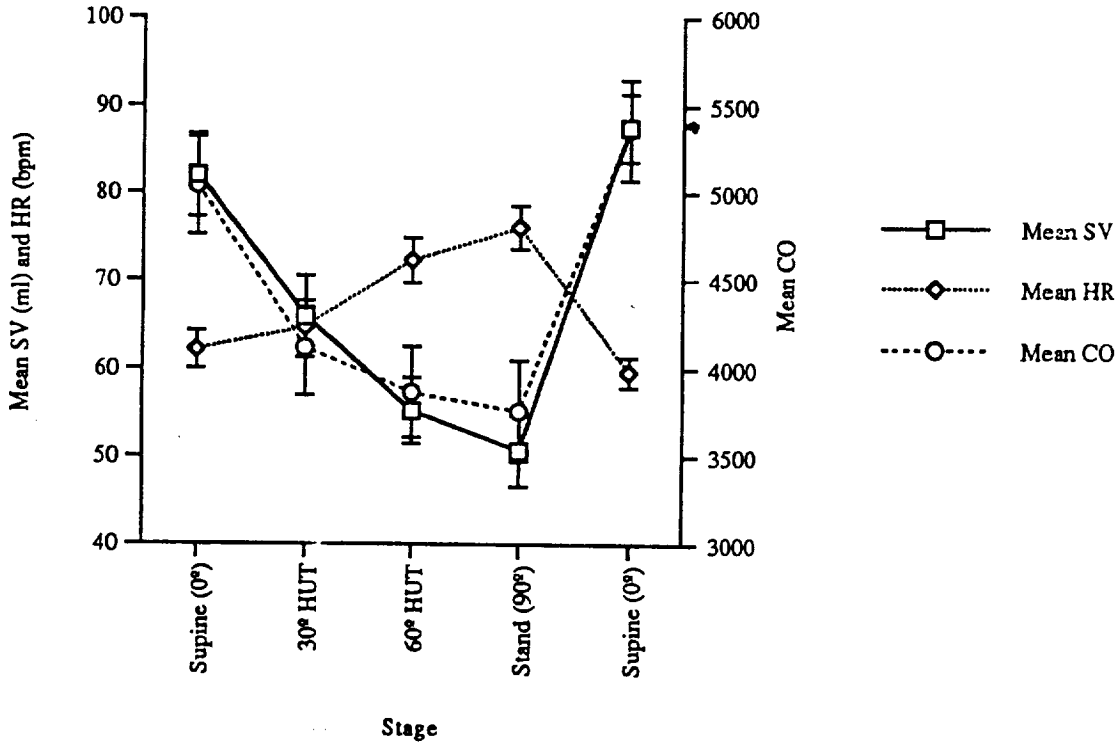


Figure 2. Mean Heart Rate(HR), Stroke Volume(SV) and Cardiac Output(CO) For All Positions

The supine heart rate was significantly different from the 30°, 60° , 90° and supine recovery positions; different from the 30° position to the 60° , 90° head-up tilt and supine recovery position; between 60° and 90° head-up tilt position and supine recovery position; and between 90° head-up tilt position and supine recovery position.

VENTRICULAR EJECTION TIME AND EJECTION VELOCITY INDEX

The mean ventricular ejection time (VET) had only slight variation during the test with the exception of the supine position of recovery at the end of the test. The mean ejection velocity index (EVI) progressively decreased from the supine baseline position to the 60° head-up tilt position. It remained essentially the same for the 90° head-up tilt position and then increased with the supine position during recovery at the end of the test.

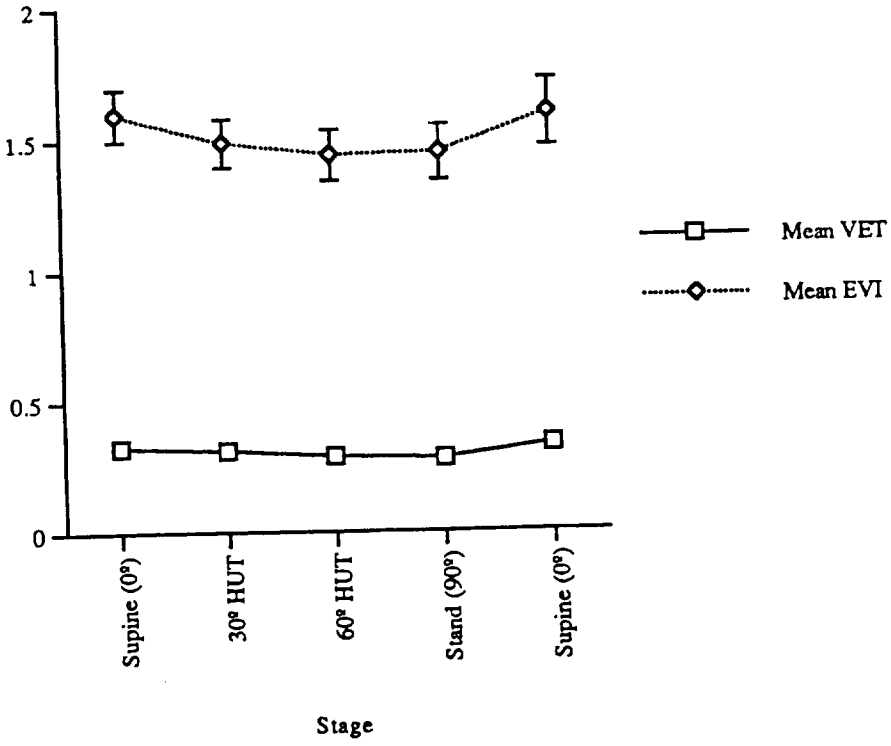


Figure 3. Mean Ventricular Ejection Time(VET) and Ejection Velocity(EVI) For All Positions

The VET was significant at the .05 level at all points: baseline supine compared with 30°, 60° and 90° head-up tilt, supine recovery position; 30° head-up tilt compared with 60° , 90° head-up tilt and supine position of recovery; between 60° and 90° head-up tilt and supine recovery position; and between 90° head-up tilt and supine recovery position. The EVI had significant differences between the supine baseline position and 60° and 90° head-up tilt positions; between 30° head-up tilt and supine recovery position; between 60° and 90° head-up tilt and supine recovery position.

The supine baseline stroke volume was significantly different from the 30° , 60° and 90° head-up tilt positions and supine recovery position; and from the 30° position to the 60° , 90° head-up tilt positions and supine recovery positions; and between 60° and 90° head-up tilt positions to the supine recovery position.

SYSTOLIC BLOOD PRESSURE

The systolic blood pressure had small incremental increases from the baseline supine position until the 90° head up tilt position when there was a slight decrease. The mean increased during the supine recovery period at the end of the test, above the baseline value. Systolic blood pressure was not significant at the .05 level in any of the comparisons of positions.

DIASTOLIC BLOOD PRESSURE

The diastolic blood pressure increased from the supine baseline position to the 60° head-up tilt position, decreased slightly at the 90° head-up tilt position and then decreased during the supine position for recovery at the end of the test.

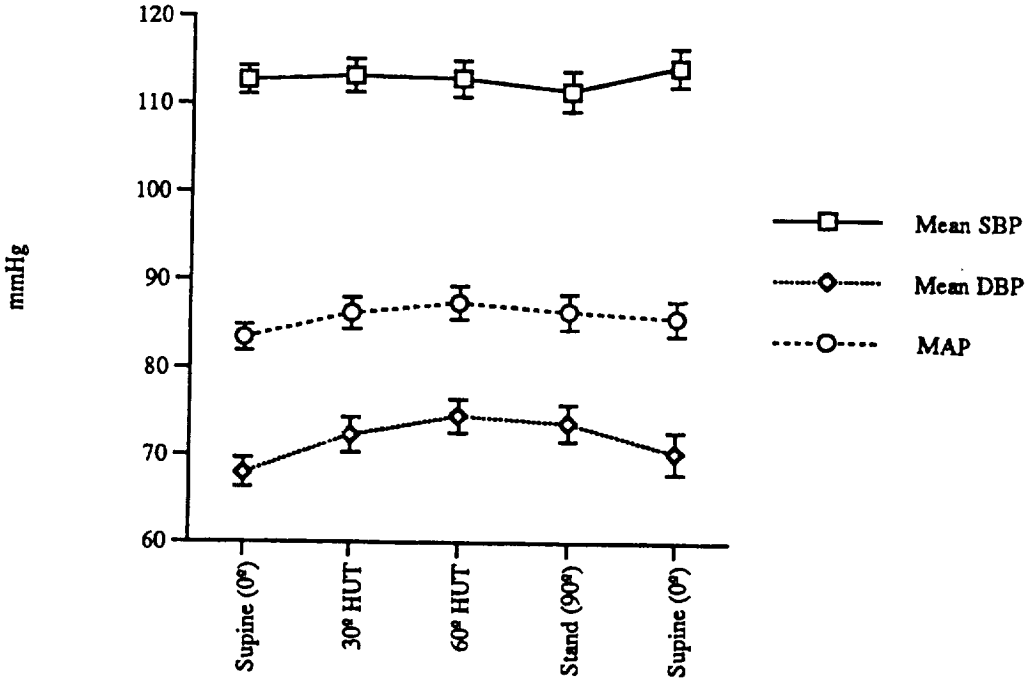


Figure 4. Mean Systolic(SBP), Diastolic Blood Pressure(DBP) and Mean Arterial Pressure(MAP) For All Positions.

The diastolic blood pressure significantly differed from the baseline supine position to the 30° head-up tilt; from supine to 60° head-up tilt position; from the supine to the 90° head-up tilt; and between 30° head-up and tilt 60° head-up tilt and supine recovery position; between 60° and 90° head-up tilt and supine recovery position.

ARTERIAL PRESSURE

The mean arterial pressure follows the diastolic blood pressure response with at greater difference at the supine recovery position.

The mean arterial pressure was significantly different from the supine baseline position to the 30° head-up tilt; from the supine to the 60° and 90° head-up tilt positions; between the 60° and 90° head-up tilt positions and supine position during recovery.

PULSE PRESSURE

The mean pulse pressure progressively decreased from the supine resting position to its minimum value at 90° head-up tilt. Its maximum value was at the supine position during recovery, but only slightly more than baseline resting.

Significance of the pulse pressure differences were from the supine baseline position to 30°, 60° , and 90° head-up tilt positions; between 30° , 60° and 90° head-up tilt positions to supine position during recovery.

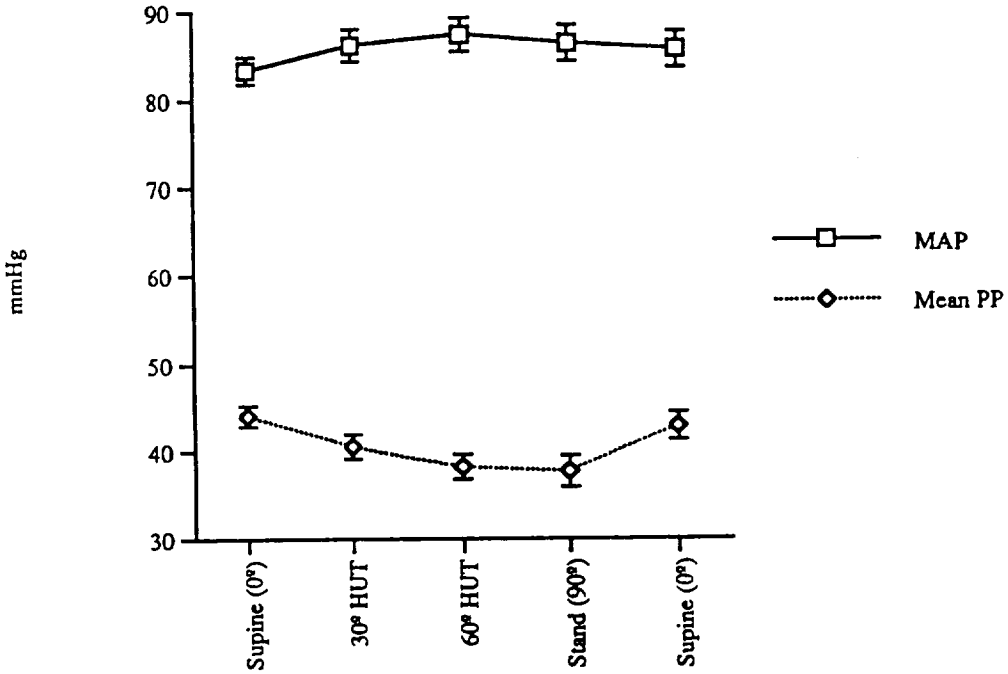


Figure 5. Mean Arterial Pressure(MAP) and Pulse Pressure(MPP) For All Positions

THORACIC FLUID INDEX

The Thoracic Fluid Index (TFI) increased in a steep slope from the supine resting position to 90 head-up tilt where it plummeted during the supine position recovery phase to a value slightly lower than supine baseline.

Significance of the TFI across tilt positions occurs at the .05 level for all comparisons except when supine resting was compared to supine recovery. As the values decrease, the fluid shifting increases.

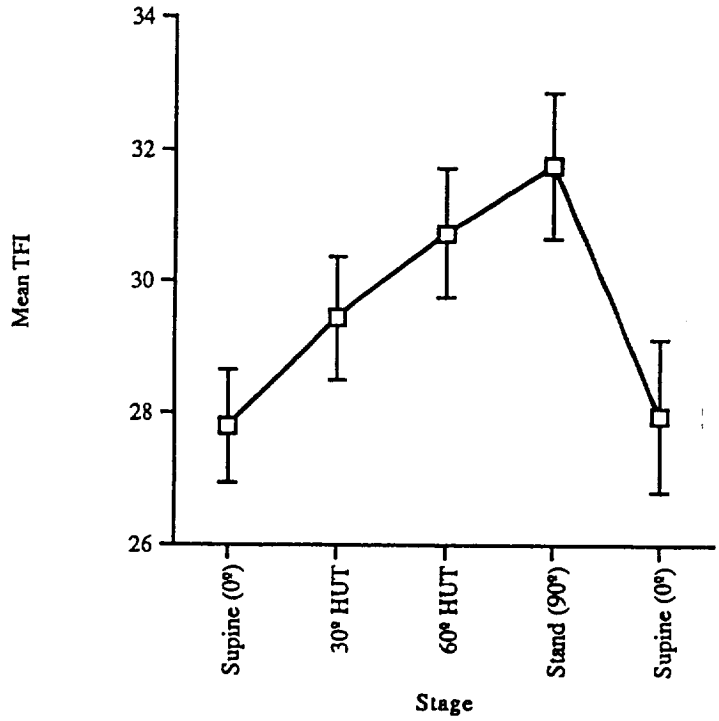


Figure 6. Mean Thoracic Fluid(TFI) Index Response For All Positions

SKIN BLOOD PERFUSION

Laser doppler monitoring of blood perfusion of the skin is reported for velocity, volume, and combined mass of blood cells taken from blood flow of the left forearm and the lateral side of the calf of the left leg. Within the skin of the arm, the mean velocity was at minimum values at the supine baseline and at the 90 head-up tilt position. During the supine position during recovery, the blood velocity values are at its maximum. The blood volume increased slightly from the supine baseline to the 90 head-up tilt and then decreased during slightly during the supine position of recovery. The combined mass of blood cells followed this same pattern as would be expected.

There was significant change in blood velocity in the skin of the arm when supine baseline position and the 90 head-up tilt position were compared to the supine recovery position. The combined mass of blood cells was significant when the supine baseline was compared with the 90 head-up tilt position; when 30 head-up tilt was compared with the 90 head-up tilt; when 30 head-up tilt was compared with supine position during recovery; and when 60 head-up tilt position was compared to 90 head-up tilt position. The blood volume was significantly different when the supine baseline was compared with 90 head-up tilt position and the supine position during recovery; 30 head-up tilt was significantly different from 90 head-up tilt and supine position during recovery; and 60 head-up tilt was significantly different from the supine position during recovery.

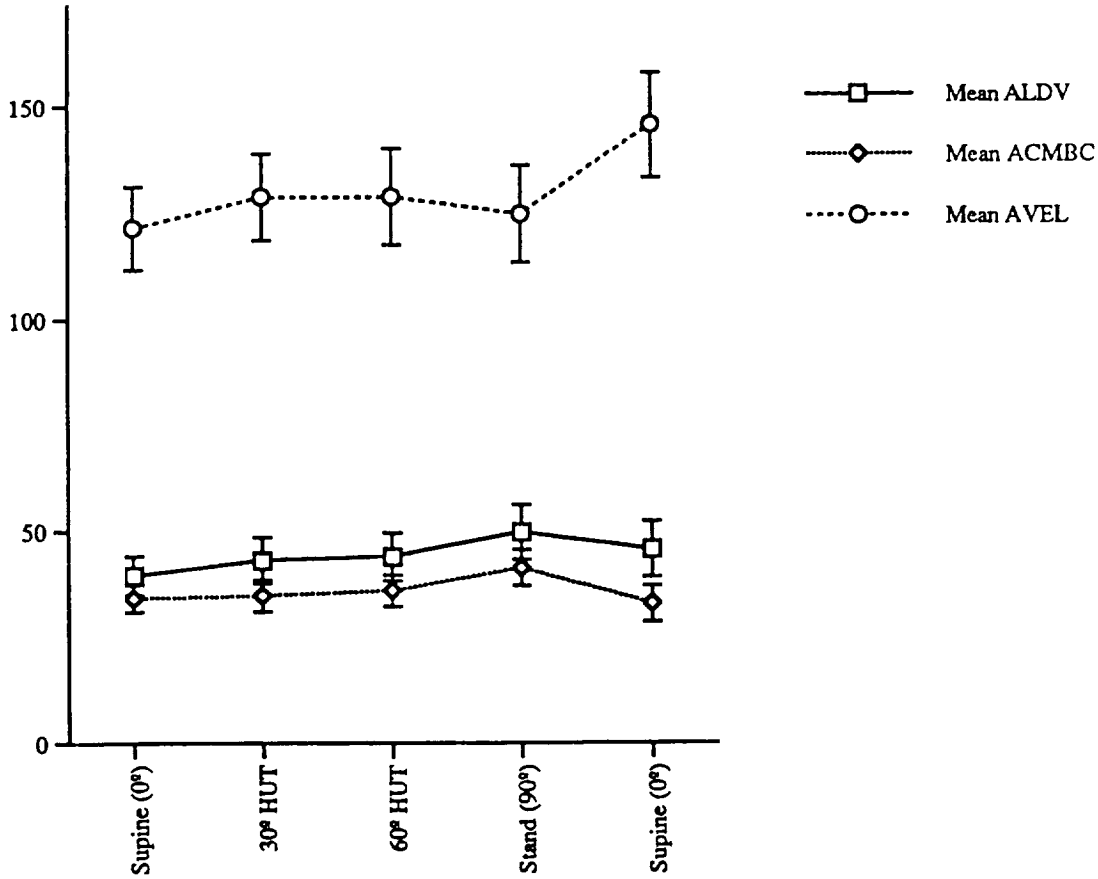


Figure 7. Mean Blood Velocity(AVEL), Volume(ALDV), and Combined Mass of Blood Cells(ACMBC) Of The Skin Of The Arm

The blood velocity of the skin of the left leg decreased progressively from the supine baseline position to 90 head-up tilt and then sharply increased during the supine position of recovery at the end of the test. The blood volume and combined mass of blood cells followed this same pattern.

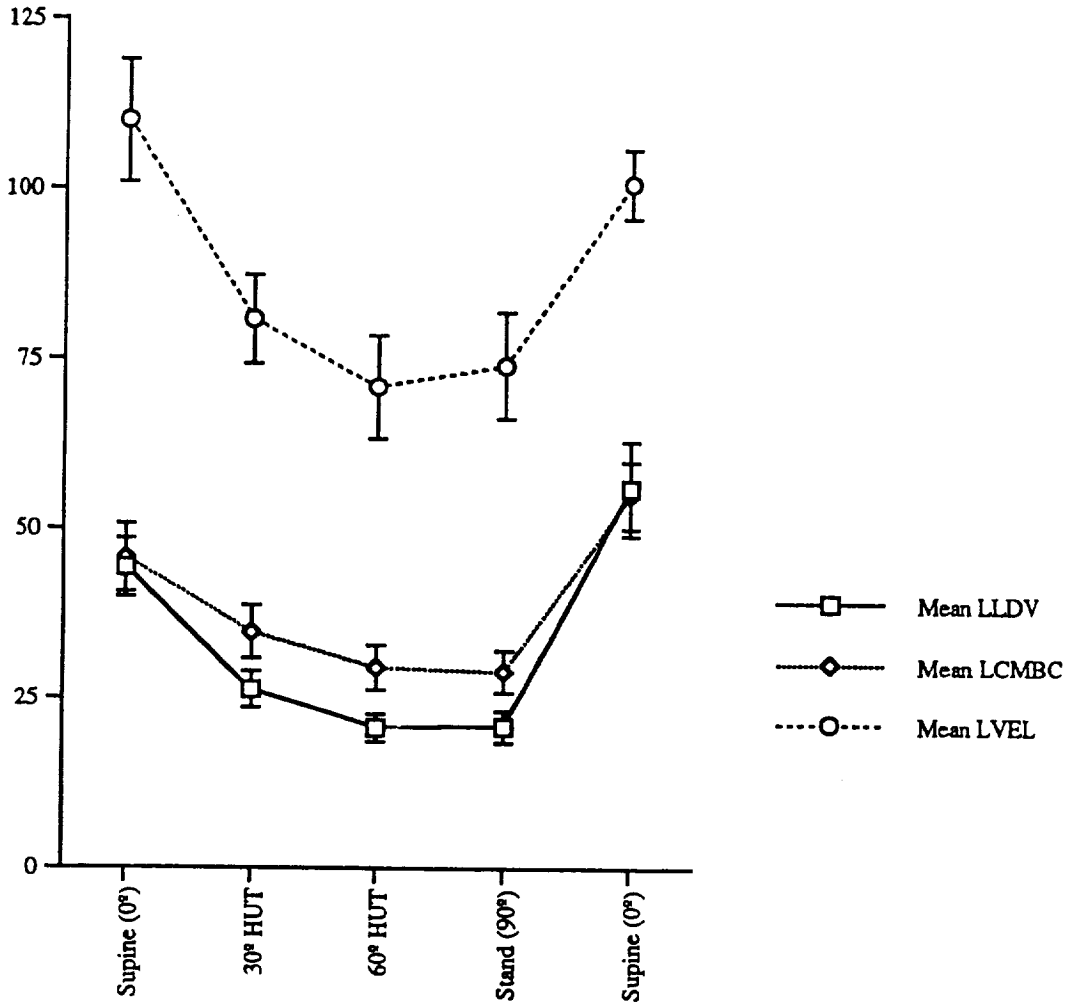


Figure 8. Mean Blood Velocity(LVEL), Volume(LLDV), and Combined Mass of Blood Cells(LCMBC) Of The Skin Of The Leg For All Positions

There were significant differences in velocity between supine baseline and 30° head-up tilt, 60° head-up tilt, and 90° head-up tilt; between 30°, 60°, and 90° head-up tilt and supine position of recovery. There were significant differences in volume from the supine baseline position and the 30° head-up tilt, 60° head-up tilt, and 90° head-up tilt; between 30° head-up tilt and supine position of recovery; and between 60° and 90° head-up tilt positions and supine position of recovery.

The areas of significance for combined mass of blood cells was from the supine baseline position to 30°, 60° and 90° head-up tilt positions. There were significant differences between the 30° head-up tilt and 60° and 90° head-up tilt and supine recovery position. The head-up tilt positions of 30°, 60°, and 90° all differed from the supine position of recovery.

DISCUSSION

Most subjects tolerated the tests very well. Only one subject indicated that a

feeling of lightheadedness or dizzyness accompanied the changes in posture from supine to the three different angles. However, each of the subjects perceived the supine recovery position as a head-down tilt when it was actually at zero position. Subjects were asked not to actively stand on the foot board at 90° but to remain as passive as possible to avoid activating skeletal muscles which would influence blood flow in the legs. Two restraining fabric bands across the body, one above the knees and the other just below the waist, helped the subjects to maintain as passive a position as possible. Even though, the subjects were in a 90° position, some perceived that their bodies were tilted beyond this angle.

All subjects showed an immediate and incrementally increasing heart rate when their posture was changed to the three head-up tilt positions. Total peripheral resistance followed this same pattern. This was accompanied by decreasing stroke volume and cardiac output. While systolic blood pressure had only slight variations until it decreased at the 90° head-up tilt, diastolic blood pressure increased incrementally with the angle changes and then also decreased at 90° head-up tilt. The mean arterial pressure followed this same response. Pulse pressure incrementally decreased with the lowest point at 90° head-up tilt. Ventricular ejection time and ejection velocity decreased only slightly from supine to 90° head-up tilt.

Mekjaviv et al (1987) induced significant ($P = .005$) increases in diastolic pressure, but no significant differences in systolic pressure when taking subjects from a supine position to a 70° head-up tilt. They made the assumption that the upright posture elevated the heart rate. Ten Harkel et al (1992) postulated that increases in diastolic and mean blood pressure and heart rate following a change in posture from 6° head-down tilt to standing reflected a redistribution of blood volume followed by vasoconstriction, a decrease in stroke volume and cardiac output.

Skin blood perfusion, measured with laser doppler, of the arm and leg differ. The velocity of blood flow of the skin of the arm increased until 90° head-up tilt, then slightly decreased. Blood volume of the skin of the arm increased to 30° head-up tilt, stablized at 60° and dropped at 90° head-up tilt. These responses reflect the activity of the baroreceptors as they cause a reflex vasoconstriction to maintain blood pressure during the standing position. The velocity of blood perfusion in the skin of the leg decreased to 60° head-up tilt, then increased at 90° head-up tilt. Blood volume in the skin of the leg decreased until 60° head-up tilt, then increased at 90° head-up tilt. The decrease in blood velocity and volume in the skin of the leg probably reflects the shift of body fluids to the chest and then at the 90° head-up tilt or standing, the blood responds to the gravitational pull and increases in velocity and volume.

Johnson (1986) demonstrated that skin blood flow is determined by more than thermoregulation and that baroreceptor reflexes have a major function in cutaneous circulation. When blood pressure regulation is challenged, cutaneous vasoconstriction occurs in normothermic subjects.

Blomquist et al (1983) estimated that changing posture from supine to an upright position increases the venous volume of the legs by approximately 650 ml. with an

additional volume of 250 ml. of blood transferred to the veins in the buttocks and the pelvic veins.

Most of the previous studies(Sander-Jensen et al, 1986; Mekjavic and Mittleman, 1987; and Ten Harkel et al, 1992) that have compared either the supine position or a head-down position ranging from -6° to -10° with responses of subjects at various head-up tilt positions have had similar responses as demonstrated in this study. However, there were differences of methodology. In most other studies, rest periods were of longer duration at the beginning of the test and between the increasing of the angles of head-up tilt. Others took the subject from the supine position immediately to a 70° head-up tilt. In this study, with a relatively short rest period of 30 minutes for baseline, significant cardiovascular changes occurred immediately upon changing posture to each increment of 30°, 60°and 90° head-up tilts within a 5 minute period of each other. While the subject was not retained in a particular position until stabilization, the responses continued to change with each angle or change in gravitational force acting upon the body. Without exception, when the subjects were returned to the supine position during recovery, all variables responded in the direction of baseline values and in some cases overshooting the baseline. Data collection was limited to a 5 minute recovery period, actual recovery may have extended beyond that time.

CONCLUSION

Based on the significance of the findings of this study, this model for assessing orthostatic responses does provide an adequate challenge to the blood pressure regulatory systems. While individual subjects may use different adaptations to incremental changes in gravity, such as increased heart rate, increased peripheral resistance, increased diastolic blood pressure or increased stroke volume; in the aggregate, the subjects demonstrated significant adaptive cardiovascular changes to orthostatic challenges which were presented to them.

REFERENCES

Blomqvist, C. G., and H. L. Stone. Cardiovascular adjustments to gravitational stress. In: Handbook of Physiology, The Cardiovascular System. Washington, DC: Am. Physiol. Soc., 1983, sect. 2, vol. III, pt. 2, chapt. 28, p. 1031.

Bungo, M. W. The cardiopulmonary system. In Space Physiology and Medicine, Ed., A. E. Nicogassian, Philadelphia: Lea & Febiger, 1989, pp. 185-191.

Fischer, J., P. M. Parker, and W. W. Shaw. Waveform analysis applied to laser doppler flowmetry. In Waveform Analysis of Laser Doppler Flowmetry, Alan R. Liss, Inc., 1986.

Greenleaf, J. E., C. E. Wade, and G. Leftheriotis. Orthostatic responses following 30-day bed rest deconditioning with isotonic and isokinetic exercise training. Aviation, Space, and Environmental Medicine, June, 1989, p. 537-542.

Johnson, J. M., G. L. Brengelmann, J. R. S. Hales, P. M. Vanhoutte, and C. B. Wenger. Regulation of the cutaneous circulation. Federation Proceedings, vol. 45, no. 13, December 1986, p. 2841-2850.

Mekjavic, I. B., C. A. Gaul, M. D. White and K. D. Mittleman. Cardiovascular responses during 70° head-up tilt: The effect of elevated body temperature and high alcohol blood levels. The Physiologist, vol. 30, no. 1, suppl., 1987, p. S56-S57.

Melchior, F. M. and S. M. Fortney. Orthostatic intolerance during a 13-day bed rest does not result from increased leg compliance. Journal of Applied Physiology, 1993, vol. 74, no. 1, p. 286-292.

Sander-Jensen, K., N. H. Secher, A. Astrup, N. J. Christensen, J. Giese, T. W. Schwartz, J. Warberg, and P. Bie. Hypotension induced by passive head-up tilt: endocrine and circulatory mechanisms. In Bradycardia During Hypotension, Washington, DC: American Physiological Society, 1986, p. R742-R748.

Ten Harkel, A.D.J., L. Beck and J. M. Karemaker. Influence of posture and prolonged head-down tilt on cardiovascular reflexes. Acta Physiol Scand, 1992, vol. 144, S604, p. 77-87.

N94-25377

DIGITAL DATA, COMPOSITE VIDEO MULTIPLEXER AND DEMULTIPLEXER BOARDS FOR AN IBM PC/AT COMPATIBLE COMPUTER

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Dean Lance Smith, Ph.D.
Academic Rank:	Associate Professor
University & Department:	Memphis State University Dept. of Electrical Engineering Memphis, Tennessee 38152
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Life Sciences Project
Branch:	Science Operations
JSC Colleague:	Walter D. Hanby
Date Submitted:	August 13, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

Work continued on the design of two IBM PC/AT compatible computer interface boards. The boards will permit digital data to be transmitted over a composite video channel from the Orbiter. One board combines data with a composite video signal. The other board strips the data from the video signal.

INTRODUCTION

The design of both boards was started towards the end of the last summer faculty fellowship period and continued under a grant during the academic year. A survey of techniques for multiplexing data with video signals was completed during the summer of 1992 [1]. Most of the specifications for the boards were completed during the academic year grant period [2]. Most of the rest of the design for both boards was completed during the current summer faculty fellowship period, and some of the schematic diagrams were drawn.

Most of the digital logic circuits are 74HC family members. Most integrated circuits can be supplied by two or more manufacturers. Some special function circuits are available from only one source. Alternatives that perform similar functions, but may not be plug-in replacement parts, are discussed in the design documentation.

Some of the designs' features and the documentation that supports some of the design decisions are discussed in more detail below.

DESIGN FEATURES AND DOCUMENTATION

Error Detection and Correction Code

A (2,1,6) convolution code with 171 (octal) and 133 (octal) generator polynomials was selected. The code has a constraint length of 7 and a coding rate of 1/2. Viterbi decoding will be used at the receiver.

The code has the power of the 1/2 rate NASA Interplanetary Standard code used on TDRSS and is supported by commercially available encoder and decoder chips. Punctured versions of the code can be used to reduce the coding (and the error detection and correction) rate.

Several codes that are supported by commercially available encoder and decoder chips were considered. Several hardware and software techniques, besides using commercially available encoder and decoder chips, were considered [3]. A paper will be submitted for publication based on a study of coding techniques and commercially available encoding and decoding chips.

Board Control Processor

The MC68HC16 was selected as the board controller for both the transmitter and receiver boards. A 68HC16Z1EVB single-board development system was ordered to assist in developing the hardware and software for the interface boards. Several microprocessor and microcontrollers were considered, including a dedicated processor constructed from HC family SSI and MSI logic chips [4]. Several of the processors studied, including the DSP56001, MC68HC000, 80C186, and 80C188, could move data quickly enough and address a large enough memory space. However, a system based on the 68HC16 will use less power.

Video Signal Processing

Several circuits that process the video signal are available from only one vendor. Alternative chips or techniques are described in the documentation [5] in case the single-source chips are not available in the future.

How It Works

A brief description of how the interface boards work was written [6].

Schematics

P-CAD was used to start drawing schematics for the transmitter board. Two schematics were completed, the convolutional encoder and the video processor [6]. Part of the video processor circuitry can be used for the receiver board. A conference paper will be submitted for publication about the encoder circuit.

RECOMMENDATIONS FOR FUTURE WORK

The remaining schematic diagrams need to be drawn. Parts need to be ordered so that the design can be tested on protoboards. Printed circuit boards need to be developed once the designs have been tested and verified, and the final design needs to be tested and verified on the printed circuit boards.

REFERENCES

[1] Dean Lance Smith and Walter D. Hanby, "Techniques for transmitting digital data on the Spacelab's analog video channel", *IEEE AES Systems Magazine*, vol. 8, pp. 42-50, May 1993.

[2] Dean Lance Smith, "Digital data, composite video multiplexer and demultiplexer boards for an IBM PC/AT compatible computer," Memphis State University, Memphis, TN, Report NAG 9-646-S, April 1, 1993.

[3] Dean Lance Smith, "Recommended error detection/correction code for the data on composite video system interface boards," memo to Walter D. Hanby, Revised August 9, 1993.

[4] Dean Lance Smith, "Processor for the data on video interface boards," memo to Walter D. Hanby, Revised August 9, 1993.

[5] Dean Lance Smith, "Alternative chips for the data on video interface boards," memo to Walter D. Hanby, August 13, 1993.

[6] Dean Lance Smith, "Data on video interface boards - how they work," memo to Walter D. Hanby, August 13, 1993.

EVALUATION OF BIOIMPEDANCE FOR THE MEASUREMENT OF PHYSIOLOGIC VARIABLES AS RELATED TO HEMODYNAMIC STUDIES IN SPACE FLIGHT

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Bruce C. Taylor, Ph.D.
Academic Rank:	Associate Professor
University and Department	The University of Akron Biomedical Engineering Akron, Ohio 44325-0302

NASA/JSC

Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Space and Biomedical Sciences Research Institute
JSC Colleague:	Janice M. Yelle, M.S.
Date Submitted:	August 13, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

Orthostatic intolerance, following space flight, has received substantial attention because of the possibility that it compromises astronaut safety and reduces the ability of astronauts to function at peak performance levels upon return to a 1g environment. Many pre- and post-flight studies are performed to evaluate changes in hemodynamic responses to orthostatic challenges after Shuttle missions. The purpose of this present project was to validate bioimpedance as a means to acquire stroke volume and other hemodynamic information in these studies.

In this study, ten male and 10 female subjects were subjected to simultaneous measurements of thoracic bioimpedance and Doppler ultrasonic velocimetry under supine, 10^{o} head down and 30^{o} head up conditions. Paired measurements were made during 6 periods of 5 seconds of breath holding, over a 2 minute period, for each of the three positions.

Stroke volume was calculated by three bioimpedance techniques and ultrasonic Doppler. The three bioimpedance methods included calculation of stroke volume:

1. using Kubicek's basic equation and obtaining peak dz/dt by numerical differentiation of the delta-z waveforms ex post facto. Resistivity values, taken from the literature, were 135Ω-cm for all cases.
2. using Sramek's modifications of Kubicek's equation and obtaining dz/dt by numerical differentiation of the delta-z waveforms ex post facto.
3. from the digital output of the BoMed® NCCOM3 Cardiovascular Monitor in which peak dz/dt is determined by electrical differentiation and an internal, noise-eliminating, algorithm.

Regression analysis showed no discernible relationship between any of the three bioimpedance techniques and Doppler ultrasound. R^2 values were <0.1 in all cases. Percent change in stroke index from the supine to head up or head down position showed all four methods demonstrated significant differences as a function of tilt with no significant differences between the 4 methods ($P<0.05$). However, additional analysis demonstrated that the major changes in stroke volume measurements were due, in part, to thoracic fluid shifts and subsequent changes in Thoracic Fluid Index (Basal Impedance).

The results of this study demonstrate a poor correlation between Doppler ultrasound and electrical bioimpedance measurements of stroke volume. The study also indicated that apparent changes in stroke volume were due, in part, to changes in basal impedance and not due to changes in dz/dt or ventricular ejection time. At this time, bioimpedance measurements, using the methods described in this paper, are not recommended for quantitative determinations of stroke volume and cardiac output in the US space program. Measurements of relative changes in stroke volume and cardiac output may be possible on the same subject however further analysis of the data will be required to determine statistical significance.

INTRODUCTION

The phenomenon of orthostatic hypotension is manifested as pre syncope or syncope that occurs in some individuals undergoing postural changes from a supine (or seated) to an upright position. The effect is caused by the sudden movement of blood from the upper body to the lower extremities during standing. This causes a reduction in the amount of blood flowing to the heart (reduced end-diastolic volume) which responds by reducing the force of contraction (Frank-Starling Law). This effect is normally short-lived because baroreceptors sense a reduction in blood pressure and respond, within a few seconds, by sending commands to correct the deficit. The normal response mechanism is complex but includes an increased heart rate, increased force of contraction and peripheral vasoconstriction. If these mechanisms fail, the net result is a reduction in oxygenated blood to the brain and subsequent syncope. Figure 1 shows an episode of orthostatic hypotension that occurred during a "Stand Test" in the cardiovascular laboratory. The test results were more pronounced, in this case, because of voluntary fluid depletion on the part of the subject.

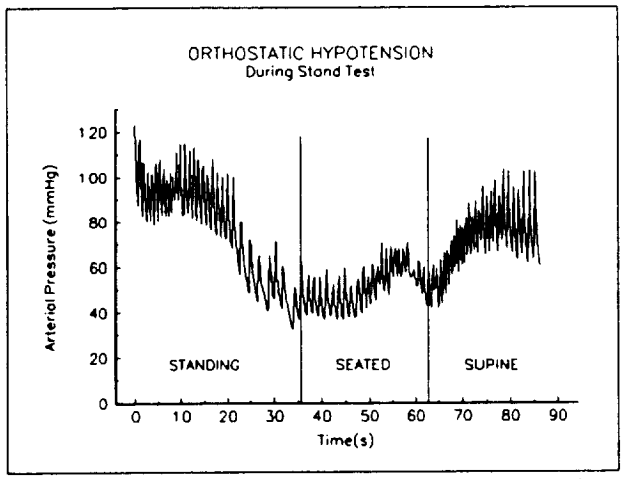


Figure 1.-An episode of orthostatic hypotension that occurred during a Stand Test. The substantial decrease in arterial pressure was corrected when the subject resumed a supine position.

In the space program this same phenomenon has been reported in a substantial number of astronauts after space flight. Microgravity exposure produces a redistribution of blood from the lower extremities. The body responds to this anomalous fluid challenge by eliminating apparent excess fluid while in microgravity. When the individual returns to the 1g environment, in a fluid-depleted-state, he is more susceptible to orthostatic hypotension. This orthostatic intolerance can cause key personnel to function below peak performance levels. Potential problems of this type are of concern and attempts are being made to eliminate the hazard.

Cardiac output is one of the most important physiologic variables to be monitored in the study orthostatic intolerance. At present non invasive cardiac output is measured, in the Cardiovascular Laboratory at JSC, using B-Mode and continuous wave Doppler ultrasonic velocimetry. The process requires measurement of the aortic outflow tract

diameter just distal to the aortic valve using pulsed echo (B-Mode) ultrasound. Beat-by-beat continuous wave (CW) Doppler measurements of aortic outflow blood velocity are then made at the suprasternal notch. Stroke volume is computed by integrating over time, the area under the ejection velocity curve to obtain mean ejection velocity for one cardiac cycle. The mean velocity times the aortic cross sectional area times the ventricular ejection time gives **stroke volume**. Stroke volume times **heart rate** gives **cardiac output**. The process of determining cardiac output using ultrasound requires sophisticated, expensive equipment and highly skilled technical support. Other non invasive methods of cardiac output measurement such as radio nuclide imaging using gated blood pool scans, magnetic resonance imaging and procedures using inhaled gases are not appropriate for use in the space Shuttle or pre- and post- landing phases of Shuttle flight.

We proposed that the measurement of stroke volume and cardiac output could be simplified through the use of electrical bioimpedance measurements. Bioimpedance cardiac monitors are less expensive, easier to apply to the test subject, require less training for proper use and are less bulky than Doppler ultrasound. Impedance cardiology, first described by Kubicek (1966) has the potential of being a major contributor to the field of non invasive cardiology except that the source of the cardiac bioimpedance signal, despite numerous attempts, has never been adequately identified. In a 1988 note on "Continuous Cardiac Output Monitoring by Electrical Bioimpedance" the American College of Cardiology concludes "Due to limitations in accuracy in certain subgroups, the use of empirical formulas, and a relative lack of correlation with other presently used non-invasive methods, the technique is considered investigational." Although numerous investigators (Geddes, et al. 1966, Judy, et al., 1969, Baker, et al., 1971, Pate, et al., 1975) have evaluated the efficacy of cardiac output by bioimpedance the space program has not accepted bioimpedance as a viable means to acquire cardiac output information.

It was the purpose of this present study to attempt to validate cardiac bioimpedance measurements using the BoMed® NCCOM3 Cardiovascular Monitor already in-house at the Johnson Space Center to compare against simultaneously-acquired ultrasonic Doppler measurements of stroke volume. Since the analog output of the NCCOM3 included ΔZ it was also possible to determine stroke volume using the Kubicek, *et al.* classic equation and the modified equation proposed by Sramek (1982) and Bernstein (1986).

Development of the Bioimpedance Equations used in this Study

Impedance plethysmography studies(Nyboer, 1970) have demonstrated that, for a volume conductor, a change in volume, ΔV, is related to a changes in electrical impedance, ΔZ, by the following equation

$$\Delta V = -\frac{\rho L^2}{Z_0^2}\Delta Z \qquad (1)$$

where

ρ is the electrical resistivity of blood at the excitation frequency used;

L is the distance between voltage sensing electrodes;
Z_0 is the mean basal impedance between voltage sensing electrodes.

If ΔZ is taken as the extrapolated maximum impedance change during systole(Ventricular Ejection Time) then ΔV represents the volume change during systole or stroke volume(SV). Kubicek determined that the peak value of the first time derivative of the ΔZ waveform, dz/dt, could be used instead of the extrapolated value. Figure 2 demonstrates this concept.

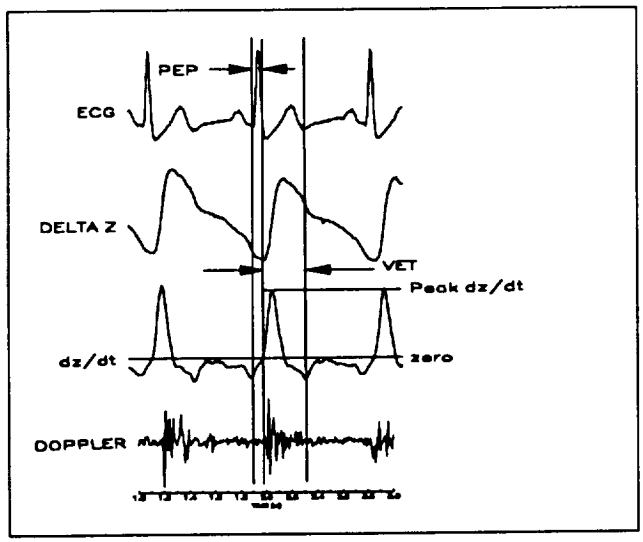


Figure 2.-An example of ECG, ΔZ, dz/dt and CW Doppler ultrasound signals taken simultaneously from the same subject. The diagram clearly demonstrates that the positive dz/dt corresponds to the ejection phase and that the peak negative dz/dt indicates the end of systole. The difference, in seconds, between these two events is the Ventricular Ejection Time (VET).

Kubicek's modification of equation (1) for stroke volume(SV_K) then becomes

$$SV_K = \frac{\left[\frac{dz}{dt}\right]_{peak} VET \rho L^2}{Z_0^2} \qquad (2)$$

where peak dz/dt and VET are taken from the analog signal as described in Figure 2.

Quail, *et al.* (1981) rearranged the Kubicek equation to solve for ρ such that

$$\rho = \frac{SV \cdot Z_0^2}{L^2\left[\frac{dz}{dt}\right]_{peak}} \qquad (3)$$

Using dogs as test subjects, by varying hematocrits and other physiological parameters and using an electromagnetic flowmeter to measure stroke volume, Quail and his associates determined that ρ was essentially constant at 135Ω-cm. Given this observation, Bernstein (1986) rearranged the Kubicek equation (2) to eliminate ρ. Since

$Z = \frac{\rho L}{\pi r^2}$ solving for ρ such that $\rho = \frac{Z \pi r^2}{L}$ and substituting in equation (2) the modified Kubicek equation becomes

$$SV = \frac{\left[\frac{dz}{dt}\right]_{peak} VET \cdot L \cdot \pi r^2}{Z_0} \qquad (4)$$

Sramek determined that the $\pi r^2 L$ represented a "Volume of Electrically Participating Tissue" (VEPT) a truncated cone whose volume was determined, experimentally, to be 1/3 that of a cylinder of the same radius and that the distance, L, between voltage sensing electrodes was 0.17*Height. Bernstein provided a nomogram to assist in the calculation of VEPT from the subject's height and weight which took into account the differences between male and female subjects. The Sramek (SV_S) equation, modified from (4) above was

$$SV_S = \frac{\left[\frac{dz}{dt}\right]_{peak} \cdot VET \cdot VEPT}{Z_0} \qquad (5)$$

Equation (5) is the same equation used by the BoMed NCCOM3 Cardiac Monitor, SV_{BM}. The difference between SV_S and SV_{BM} lies in the method that the peak dz/dt is obtained. This is explained further in the Methods section.

MATERIALS AND METHODS

Ten normal female and 10 normal male volunteers, who had current NASA Class III physical examinations, were selected for this study. The subjects read and signed the NASA Human Research Minimal Risk Informed Consent Form prior to testing. Each subject drank 0.5 Liters of water just prior to the test in order to assure a reasonable level of hydration. The height and weight of each subject was recorded and entered into the BoMed® NCCOM3® Cardiovascular Monitor (BoMed Medical Manufacturing, Ltd., Irvine, CA 92718). Each subject was fitted with BoMed #04-030 Pre-gelled disposable electrodes . Two electrode pairs were placed on the neck with the uppermost member of each pair(current electrodes) even with the angle of the jaw. Two lower pairs of electrodes were placed on each side of the thorax with the uppermost electrode of each pair at the level of the ziphoid process.

The vertical distance, L, between the two innermost (sensing) electrodes was recorded for use in calculating stroke volume by Kubicek's method. Circumferential band electrodes were not used in this study .

Ultrasound measurements were made on a Biosound Genesis II®(Biosound Corporation, Indianapolis, IN 46250) echocardiograph system with a 2.0MHz CW Doppler probe. Two dimensional images of the Aortic Outflow tract (Ao2D) were also made with the Genesis II operating in the B-Mode.

Data for each subject were recorded on a TEAC (Teac, America Inc., Montebello, CA 90640) MR-30 Cassette Data Recorder running at 9.52 cm/s. This speed provided bandwidths of 100Hz - 18.8kHz in the direct-record mode used for the CW Doppler signals and DC - 2.5kHz in the FM mode.

Output signals from the NCCOM3 include the following signals: ΔZ(1V = 0.1Ω), dz/dt(1V = 1Ω/s), ECG(1v = 1mV) and Digital Signals(RS-232 9600 Baud, 11 bit format, 1 Start bit, 7 data bits, even parity, 2 stop bits)

In the slow mode, the NCCOM3 averages 16 beats and outputs the following derived parameters via the serial I/O port: Cardiac Output (L/m), Stroke Volume(ml), End Diastolic Volume(ml), Peak Flow(ml/s), Ejection Fraction(%), Heart Rate(beats/m), Thoracic Fluid Index (Z_0), Index of Contractility(sec^{-1}), Ejection Ratio(%), Systolic Time Ratio(%), Acceleration Index(sec^{-2}), Time (every minute)

In the fast mode, the NCCOM3 provides the following data on a beat by beat basis: Cardiac Output(L/m), Stroke Volume(ml), End Diastolic Volume(ml), Peak Flow(ml/s), Ejection Fraction(%), Heart Rate(beats/m), Time(every minute)

Data stored on magnetic tape included the following: CW Doppler audio from the echocardiograph(direct record mode); ECG from the NCCOM3; ΔZ from the NCCOM3; dz/dt from the NCCOM3; Event Marker; Serial (RS-232, 2400 Baud) Baud rate shifted data from the NCCOM3 and Voice

Video and audio output from the Genesis II echocardiograph were also recorded on a Panasonic 6300 Video Cassette Recorder for post processing of Ao2D and CW Doppler signals. Serial data from the NCCOM3 were at 9600 Baud which required a bandwidth beyond the capability of the TEAC MR-30 recorder. Baud rates were shifted to 2400 Baud through the use of a Black Box® (Black Box Corporation, P.O. Box 12800, Pittsburgh, PA 15241-0800) CMA02A interface adapter. The CMA02A buffers the incoming 9600 Baud data and outputs it at 2400 Baud. The 2400 Baud signals are then recorded on the TEAC Recorder on an FM channel.

After the subject was fully instrumented, appropriate data entered into the NCCOM3 and recorded on tape, the Ao2D measurements were taken with the Genesis II Echocardiograph while the subject was placed left-side-down. After the aortic dimensions

were measured, the subject was required to lie in the supine position on a Stryker Circle® Bed (Stryker Corporation, Kalamazoo, MI 49001).

During the next four minutes of testing, the following protocol was followed for each subject at each of three (supine, 10^{o} head down and 30^{o} head up) positions.

Minute 1
NCCOM3 set to the Slow Mode (averages every 16 "good" beats) and CW Doppler measurements made at the Suprasternal Notch.
Minutes 2 and 3
The NCCOM3 was switched to the Fast Mode (beat x beat analysis of stroke volume and cardiac output) and CW Doppler signals taken.
Subject was required to breath normally for 15 seconds then hold breath for 5 seconds. This process was repeated 6 times over the next two minutes.
Minute 4
One minute period of data acquisition during which the NCCOM3 was placed in the Slow Mode and ultrasound measurements taken.

Sequencing was done through the audio channel on the VCR and TEAC tapes, event marks on the TEAC tape and Image Reversal (Flash) on the Genesis II Echocardiograph.

ECG, ΔZ and event waveforms were transcribed to PC floppy disks by playing back the FM analog tapes through a DATAQ (DATAK Instruments, Akron, Ohio) analog to digital converter using DATAQ's CODAS Signal analysis software. Each data channel was digitized at 100 samples per second. The dz/dt data were not digitized because dz/dt could be obtained by differentiating the ΔZ channel using the CODAS software. Peak values of dz/dt and VET could then be determined manually for the SV_K and SV_S and automatically, via the NCCOM3 unit for SV_{BM}

Figure 3 shows representative tracings of the transcribed data used for analysis. This protocol permitted collection of 4 successive heart beats and the associated impedance signals. For individuals with extremely low heart beats, with R-R intervals exceeding 1.25s, only 2 or 3 successive beats were acquired.

RESULTS

After data collection, the analog and serial data were transferred to PC-compatible floppy disks. Analog data were separated by event (supine, 10^{o} head down, 30^{o} head up). ECG, ΔZ and event mark were transcribed initially. Then using the CALC feature of the CODAS Signal Processing Software, the ΔZ data were differentiated to produce a dz/dt channel. Peak dz/dt(referred below as IMPSIG) and VET measured during each breath-hold sequence were then manually determined and transferred to a Quattro® (Borland International, Inc., Scotts Valley, CA 95067-0001) spread sheet for processing. Stroke volumes using Kubicek's and Sramek's formulae were then calculated using the following equations:

$SV_K = (IMPSIG/10*VET*135*L^2)/TFI^2$
$SV_S = (IMPSIG/10*VET*VEPT)/TFI$

Since TFI data were not available on a beat-to-beat basis, the TFI values were determined just before and just after the 2 minute test series. The two readings, when different, were averaged.

Stroke volumes from the BoMed NCCOM3 were taken directly from the serial output data for the same time interval as the impedance data and entered into the spread sheet as SV_{BM}. Doppler echocardiographic data were obtained from the Genesis II Echocardiograph . These data were obtained for the same beats as the impedance data. In some cases, all 4 beats were not obtainable because of frame misalignment on the VCR. In these cases data were obtained for as many beats as possible.

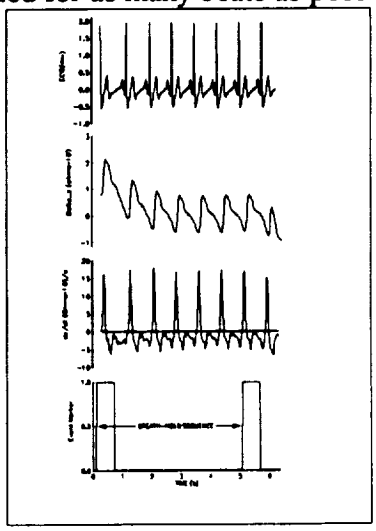


Figure 3.-Digitized data showing event marker that denotes start and finish of 5 second breath-hold sequence. Six of these sequences are performed during a 2 minute test interval at each position. The breath-hold sequence(arrows) eliminates the respiration artifact in the impedance signals.

The data were reduced by averaging the beats taken during the 5 second breath-hold sequences. In this way, each data point represents the average stroke volume, for Doppler and Impedance methods, for each 5 second period. To normalize the data for size differences between subjects, each Stroke Volume value was divided by Body Surface Area (BSA) to obtain **Stroke Index**.

An Analysis of Variance (ANOVA) was performed to test significant effects for replications, sex, and event(tilt). The results are Shown in Table I with values <0.05 considered to be statistically significant.

These results indicated there were no significant differences in replications so that the 6 data points for each condition could be combined. Also of note is the fact that there was a significant interaction for the sex-by-tilt grouping. Analysis of simple effects showed that all impedance methods were highly significant for sex($P \leq 0.001$). Doppler

was not significant for sex except for the 30° head-down tilt case which was highly significant (P≤0.001)

Table 1.-ANOVA results to test the significant effects for sex, replications(REPS) and tilt.

	SEX	REPS	TILT	SEX by REPS	SEX by TILT	TILT by REPS	3 way
SI_{BM}	.0001	.9922	.0001	.8044	.0046	.999	.999
SI_K	.0001	.9802	.0001	.98	.0922	.9917	.984
SI_S	.0001	.9878	.0001	.9813	.0026	.9878	.910
SI_D	.0001	.6037	.0001	.9181	.0132	.9919	.946

To observe the effects of tilt, the percent change in stroke index for the three tilt events, using combined data, were plotted. These results are shown in Figure 4. Statistical analysis (P≤0.05)showed that differences between events 1 and 2 were not significant but that event 3 was different from both groups 1 and 2. There were no significant differences between groups, i.e., all 4 methods produced equivalent results. Scatter plots of stroke indices of all three impedance methods are shown in Figure 5.

To test the validity of the NCCOM3 algorithm to detect and differentiate the ΔZ signal, a scatter plot of SI_S by SI_{BM} was done and shown in Figure 6.

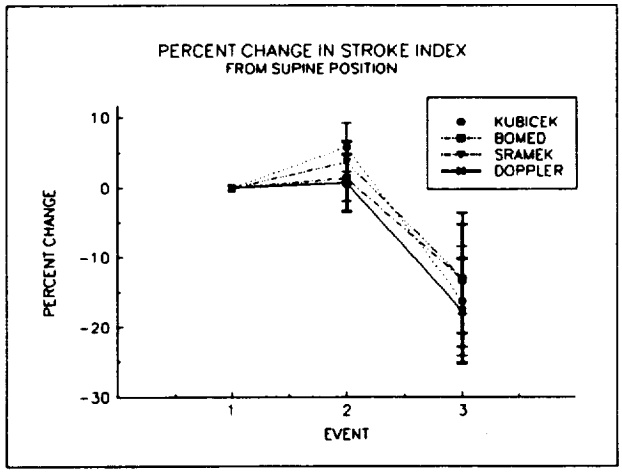


Figure 4.-Combined data showing percent change in stroke index as caused by head-up(2) and head-down(3) tilt from the supine(1) position. Vertical bars are 95% confidence limits.

DISCUSSION

In looking at the percent change (Figure 4) data there was a significant shift in the cardiac output from the 10° head-down position to the 30° head-up position. The magnitude of change from the head-down to head-up positions were roughly the same for all 4 methods. The Doppler and Sramek's method showed parallel effects. The BoMed method showed a slightly reduced sensitivity (lower slope) and the Kubicek method showed a slightly increased sensitivity(steeper slope) to positional changes. Using Duncan's Multiple Range Test to compare means there were no significant differences

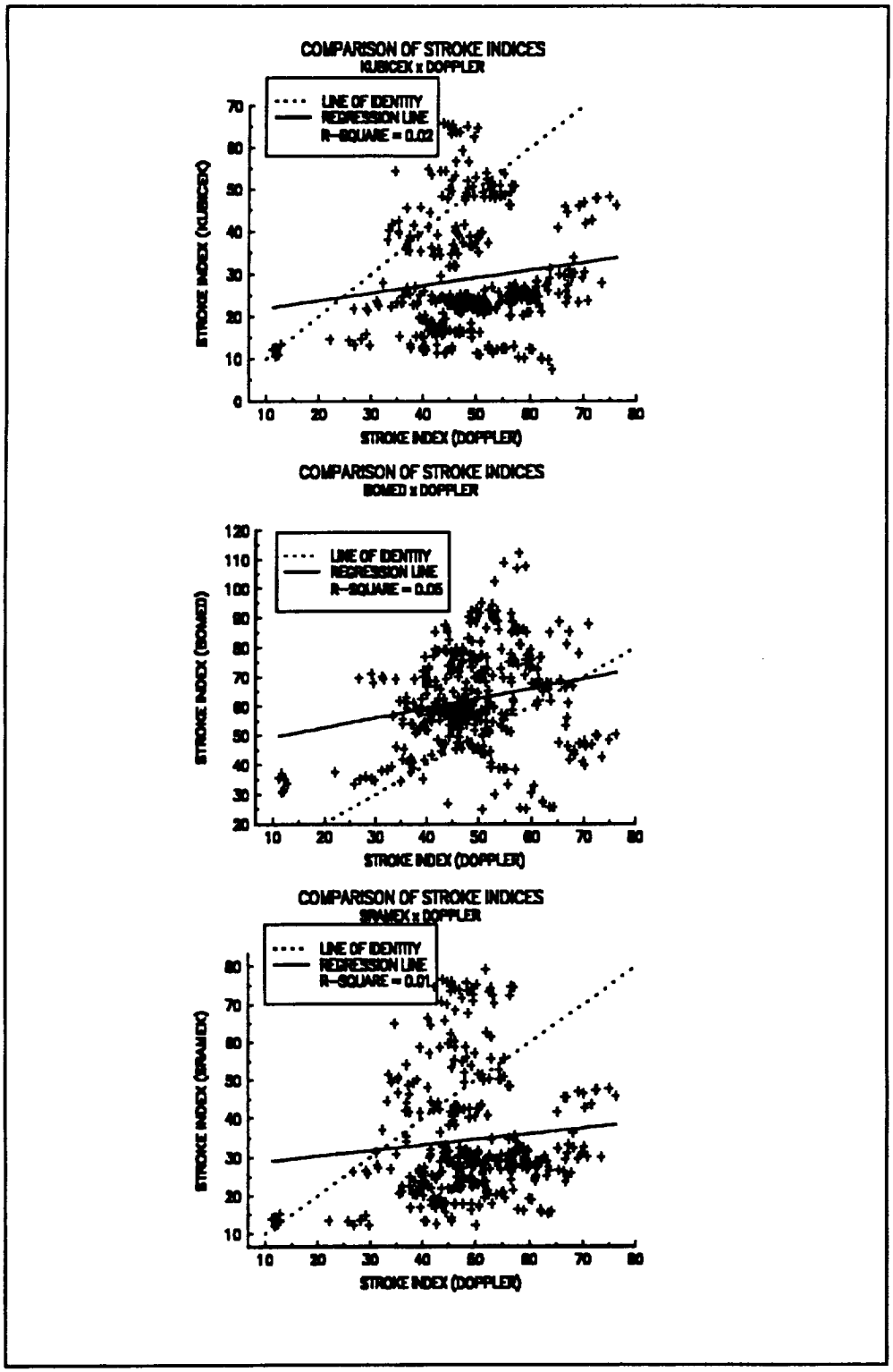


Figure 5.-Comparison of stroke indices of Kubicek x Doppler, BoMed NCCOM3 x Doppler and Sramek x Doppler. The solid line represents the linear regression line for the data. The dotted line represents the line of identity along which the data would lie if the variables were equivalent. R^2 values for the three groups are 0.02, 0.05 and 0.01 respectively.

among any of the 4 methods within any treatment group (degree of tilt) however there were significant interactions between degree of tilt and sex.

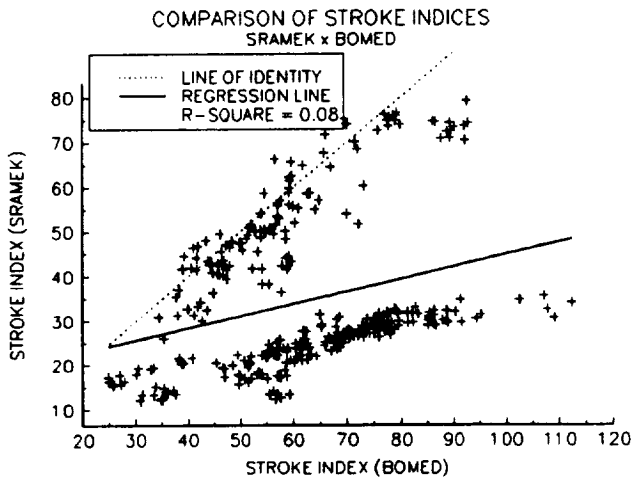


Figure 6.-Scatter plot of stroke index calculated by Sramek's equation and the BoMed NCCOM3 cardiovascular monitor. The data should fall on the line of identity with an R^2 value of 1.0. The data exhibit a bi-modal distribution with a low R^2 value of 0.08. The reason for this distribution is not clear.

Looking at the scatter plots of the three impedance methods and Doppler, there appeared to be no correlation among any of the events. Correlation coefficients were less than 0.1 in all cases. This indicates that either impedance (all three methods) or Doppler or both are not measuring stroke volume. Since the Doppler echocardiographic analysis of percent change studies produced the expected results, one can suspect that the echocardiograph is measuring a real quantity, namely stroke volume. On the other hand, one cannot assume this of the bioimpedance signals because of the changes in TFI that occur with postural changes.

One is then led to consider these two apparent contradictory results, namely acceptable results when percent change from supine is considered and random results when comparing Doppler to bioimpedance. The assumption must be that something else is affecting the impedance signals. A likely culprit is the TFI signal which also varies with positional changes. A look at equations (2) and (5) show that a change in TFI will produce changes in stroke volume in the direction observed. It should also be noted that the change in SV_K will produce a steeper slope because of the squared term that appears in the denominator.

In order to determine to what extent the TFI changes would have on stroke volume or stroke index, the data were recalculated by keeping all variables constant except for TFI. The constant was taken as the mean dz/dt, mean VET, mean L and mean VEPT over all trials. For the Kubicek equation, ρ was taken as 135Ω-cm. The results of this investigation were plotted against the percent change data and are shown in Figure 7.

This analysis demonstrates that a major portion of the change that is seen in the three test groups is due to changes in TFI and not due to measured changes in dz/dt and

VET. Statistical analyses demonstrated highly significant differences in the TFI data by sex, by tilt and by sex*tilt ($P \leq 0.001$).

A second set of statistical analyses keeping TFI constant showed the following:

For the Kubicek Data, sex was not significant but tilt was ($P \leq 0.05$) when moving from the head-down(2nd) to the head-up(3rd) position.
For the Sramek data, sex and tilt were both significant ($P \leq 0.05$) when moving from the 2nd to 3rd position.

Comparison of the stroke index as calculated from Sramek's formula and the digital output of the NCCOM3 showed no overall correlation but some of the data did appear to fall on the line of identity. The rest of the data accumulated around a second slope giving the data a bi modal appearance. At this time it is not known why this bi modal distribution occurred nor is it known which data belong to what group.

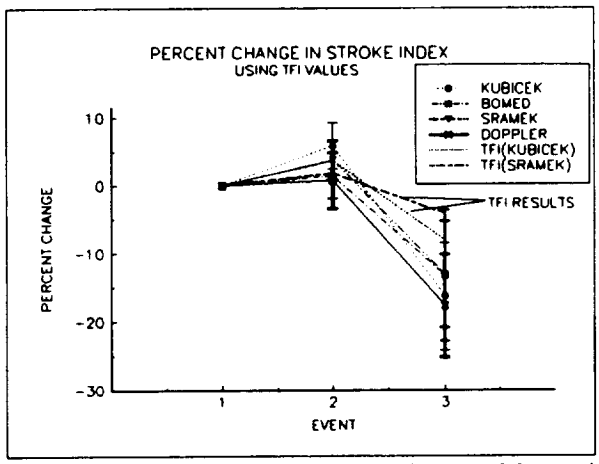


Figure 7.-Percent change in stroke index from the supine position with TFI results added. The TFI data represent changes in stroke index that can be attributed to TFI only.

CONCLUSIONS

This study was initiated in order to validate the bioimpedance measurements of stroke volume and cardiac output. Doppler echo velocimetry has been used in the cardiovascular laboratory and on several missions aboard the space Shuttle. The problem with the technique is the cost of the equipment and the skill level necessary to use it. Bioimpedance, on the other hand, has been purported to be accurate yet inexpensive and does not require the skill level of echocardiography.

This study, however, does not support the use of bioimpedance measurements in conjunction with the space program. Analysis of the data collected during this study show no correlation between stroke volume measured by echocardiography and that measured by bioimpedance. Furthermore, it appeared that the changes that were detected were due to shifts in thoracic fluid and consequential changes in Thoracic Fluid Index (Basal Impedance, Z_0).

Comparisons of the BoMed NCCOM3 monitor with manually derived values, using the same equations, showed no overall correlation though a small group of data points did appear to correlate well. This result indicates that the algorithm that is internal to the NCCOM3 does not agree well with manually-derived results and that neither of the methods is particularly quantitative.

Reasons for the apparent lack of correlation are unknown at this time. If we assume that the manual calculations of Kubicek's and Sramek's equations were faulty, then the BoMed results should have been acceptable, but they were not. If we assume that the BoMed internal algorithm was faulty, then Kubicek's and Sramek's methods should have worked out well, but they did not. Several of the subjects dz/dt waveforms were difficult to interpret because of double dz/dt peaks or because of indistinct VETs. These data were in the minority and would not affect the overall results.

It is possible that the BoMed NCCOM3 used was producing noisy ΔZ signals which would lead to noisy, randomized results. It is also possible that the BoMed electrode system used is not at all suitable for this type of study. Until these questions are answered, however, use of this system to study stroke volume and cardiac output is not recommended. Relative changes in stroke volume or cardiac output and the measurement of TFI may find some use in the program but further research needs to be conducted..

It is recommended that this series of experiments be repeated with the following changes:

Use a different bioimpedance monitor, for example use the Minnesota Impedance Cardiograph or build one.

Since the Ventricular Ejection Time is sometimes difficult to determine using the dz/dt waveform, it would be better to identify systole acoustically with a phonocardiograph microphone.

Use circumferential electrodes since it is not clear whether the BoMed electrodes provide the uniform current densities that this technique requires.

Continue to use Kubicek's and Sramek's formulae since variables for Sramek's formula are easier to obtain and do not require calculations based on the subject's hematocrit.

REFERENCES

Baker, LE, Judy, WV, Geddes, LA, *et al.* The measurement of cardiac output by means of electrical impedance. *Cardiovasc, Res. Ctr. Bull.*, 9:135-145, 1971.

Bernstein, DP A new stroke volume equation for thoracic electrical bioimpedance:Theory and rationale. *Crit. Care Med.* 14:904-909, 1986.

Geddes, LA, Hoff, HE, Mello, *et al.* Continuous measurement of ventricular stroke volume by electrical impedance. *Cardiovasc. Res. Bull.* 4:118-131, 1966.

Judy, WV, Langley, FM, McCowen, KD, *et al.* Comparitive evaluation of thoracic impedance and isotope dilution methods for measuring cardiac output. *Aerosp. Med.* 40:532-536, 1969

Kubicek, WG, Karnegis, JN, Patterson, RP, *et al.* Development and evaluation of an Impedance Cardiac Output System. *Aerospace Med.* 37:1209-1212, 1966.

Nyboer, J. *Electrical Impedance Plethysmography* (2nd ed.), Charles C. Thomas, Springfield, Il, 1970.

Pate, TD, Baker, LE and Rosborough, JP The simultaneous comparison of the electrical impedance method for measuring stroke volume and cardiac output with four other methods. *Cardiovasc. Res. Ctr. Bull.* 14:39-52, 1975.

Quail, AW, Traugott, FM, Porges, WL, *et al.* Thoracic resistivity for stroke volume calculation in impedance cardiography. *J. Appl. Physiol.* 50:191-195, 1981.

Sramek, BB Cardiac output by electrical impedance. *Med. Elect.* April:93-97, 1982

Note:

W.G. Kubicek has compiled and published *A Bibliography of Publications Related to Impedance Cardiography* which contains 546 listings on the subject. This compendium of civilian and space flight, can be obtained by writing to:

W.G. Kubicek, Ph.D., Professor Emeritus
University of Minnesota Medical School
4180 Edmund Blvd.
Minneapolis, MN 55406

Also note that numerous publications by Sramek, the developer of the competing BoMed Impedance Cardiograph, do not appear to be listed in this compendium. Sramek, tends to publish in "throw-away" magazines (Medical Electronics, Critical Care Medicine and his own publications) and it is possible that Dialog Information Services, Inc. (the source of Kubicek's bibliography) does not index these magazines.

N94-25379

Conceptual Communications System Designs in the 25.25–27.5 and 37.0–40.5 GHz Frequency Bands

Final Report
NASA/ASEE Summer Faculty Fellowship Program -1993
Johnson Space Center

Prepared By:	Michael W. Thompson, Ph. D.
Academic Rank:	Assistant Professor
University & Department:	Department of Electrical Engineering Colorado State University Fort Collins, Colorado 80523
NASA/JSC	
Directorate:	Engineering
Division:	Tracking and Communications
Branch:	Systems Design
JSC Colleague:	Shayla Davidson
Date Submitted:	August 18, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

Future space applications are likely to rely heavily on Ka-band frequencies (20-40 GHz) for communications traffic. Many space research activities are now conducted using S-band and X-band frequencies, which are becoming congested and require a degree of pre-coordination. In addition to providing relief from frequency congestion, Ka-band technologies offer potential size, weight and power savings when compared to lower frequency bands.

The use of the 37.0-37.5 and 40.0-40.5 GHz bands for future planetary missions were recently approved at the 1992 World Administrative Radio Conference (WARC-92). WARC-92 also allocated the band 25.25–27.5 GHz to the Intersatellite Service on a primary basis to accomodate Data Relay Satellite return link requirements. Intersatellite links are defined to be between artificial satellites and thus a communication link with the surface of a planetary body, such as the moon, and a relay satellite orbiting that body are not permitted in this frequency band.

This report provides information about preliminary communications system concepts for forward and return links for Earth-Mars, and Earth-Lunar using the 37.0-37.5 (return link) and 40.0-40.5 (forward link) GHz frequency bands. In this study we concentrate primarily on a conceptual system for communications between Earth and a single lunar surface terminal (LST), and between Earth and a single Mars surface terminal (MST). Due to large space losses, these links have the most stringent link requirements for an overall interplanetary system. The Earth ground station is assumed to be the Deep Space Network (DSN) using either 34 meter or 70 meter antennas.

We also develop preliminary communications concepts for a space-to-space system operating at near 26 GHz. Space-to-space applications can encompass a variety of operating conditions, and we consider several "typical" scenarios described in more detail later in this report. Among these scenarios are vehicle-to-vehicle communications, vehicle-to-geosyncronous satellite (GEO) communications, and GEO-to-GEO communications. Additional details about both the interplanetary and space-to-space communications systems are provided in an "expanded" final report which has been submitted to the Tracking and Communications Division (TCD) at the NASA Johnson Space Center.

Introduction

Future space applications are likely to rely heavily on Ka-band frequencies (20-40 GHz) for communications traffic. Many space research activities are now conducted using S-band and X-band frequencies, which are becoming congested and require a degree of pre-coordination. In addition to providing relief from frequency congestion, Ka-band technologies offer potential size, weight and power savings when compared to lower frequency bands.

The use of the 37.0-37.5 and 40.0-40.5 GHz bands for future planetary missions were recently approved at the 1992 World Administrative Radio Conference (WARC-92). WARC-92 also allocated the band 25.25–27.5 GHz to the Intersatellite Service on a primary basis to accomodate Data Relay Satellite return link requirements. Intersatellite links are defined to be between artificial satellites and thus a communication link with the surface of a planetary body, such as the moon, and a relay satellite orbiting that body are not permitted in this frequency band.

This report provides information about preliminary communications system concepts for forward and return links for Earth-Mars, and Earth-Lunar using the 37.0-37.5 (return link) and 40.0-40.5 (forward link) GHz frequency bands. Preliminary studies (see [1, 2, 3, 4]) have investigated the technical feasibility of millimeter wave frequencies for lunar and Mars exploration activities as proposed for the Space Exploration Initiative (SEI). The communications infrastructure proposed for the SEI program is illustrated in Figure 1. This figure depicts a complex communications scenario which has evolved to include elements such as point-to-point surface communications on both the moon and Mars, mobile communications using both manned and robot controlled rovers, message routing and relay satellites for nearly complete planetary coverage. Initial lunar and Mars missions are likely to be less complex with fewer links and components. In this study we concentrate primarily on a conceptual system for communications between Earth and a single lunar surface terminal (LST), and between Earth and a single Mars surface terminal (MST). Due to large space losses, these links have the most stringent link requirements for an overall interplanetary system. The Earth ground station is assumed to be the Deep Space Network (DSN) using either 34 meter or 70 meter antennas.

We also develop preliminary communications concepts for a space-to-space system operating at near 26 GHz. Figure 2 illustrates the types of applications envisioned for this frequency band. Space-to-space applications can encompass a variety of operating conditions, and we consider several "typical" scenarios described in more detail later in this report. Among these scenarios are vehicle-to-vehicle communications, vehicle-to-geosyncronous satellite (GEO) communications, and GEO-to-GEO communications.

Bit Rate Requirements

Space communications applications rely heavily on voice, command, scientific and visual information. Due to the high data rates involved, image/video requirements are usually a major factor in determining the bit rate requirements for various communications links. Link margin calculations (described later in this report) reveal that, for the near future, feasible return links from Mars will be in the 0.5–5 Mbps range. In fact, the 5 Mbps figure is derived by pushing the antenna, transmitter power, and receiver noise figure parameters to their outer reasonable limits, and thus initial return links from Mars are likely to operate at

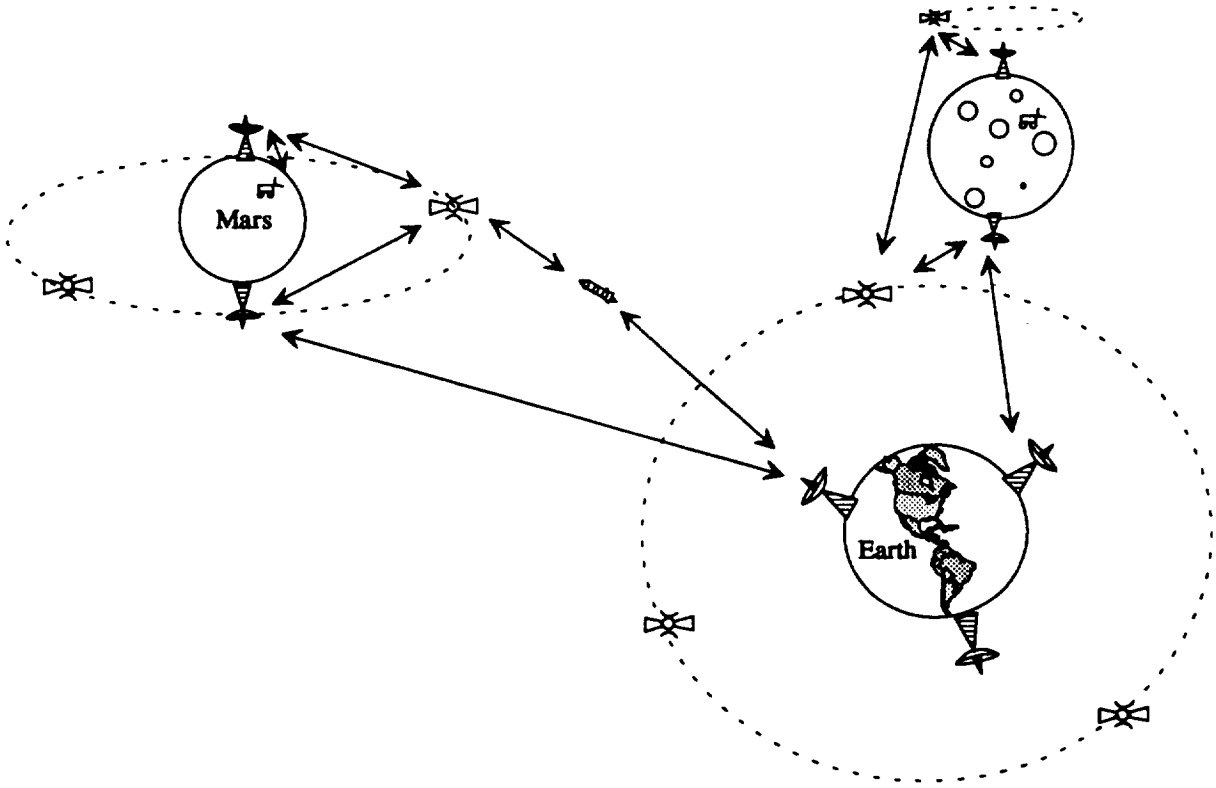


Figure 1: Telecommunications architecture for Mars and Lunar Explorations

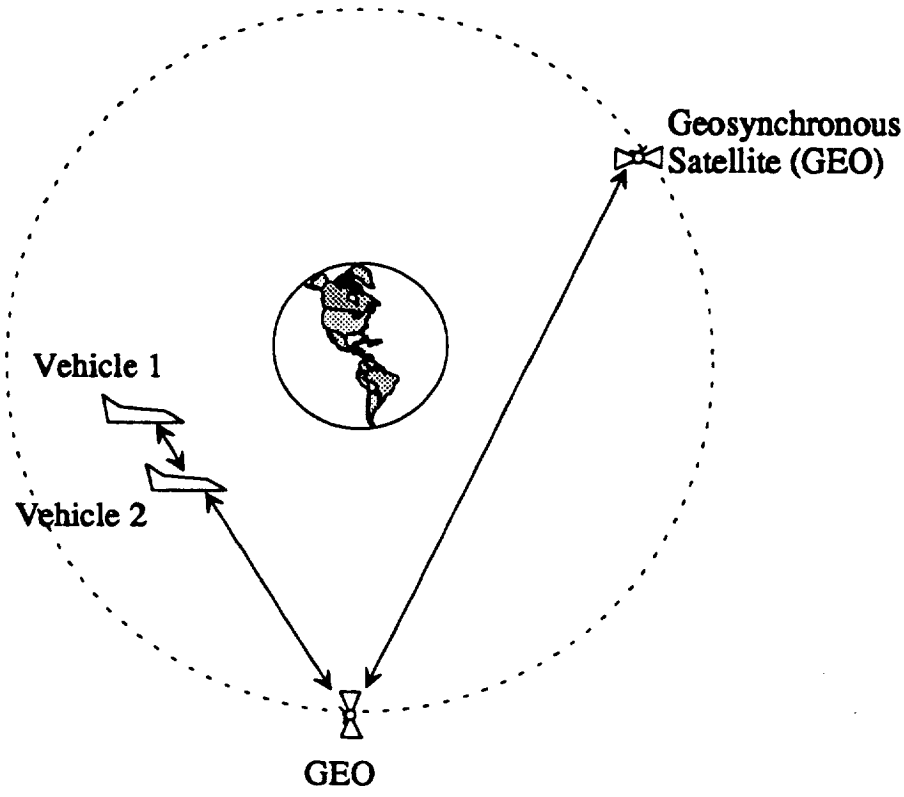


Figure 2: Space-to-Space Line Diagram

a lower maximum return rate, possibly on the order of 1 Mbps under clear sky conditions. However, the near-term development of Ka-band technology to support relatively large data rates for lunar and space-to-space applications appears to be feasible. The space loss for lunar applications is up to 60 dB less than that of Mars. Thus, lunar links are capable of supporting significantly higher data rates.

A combination of (at least) two approaches may be used to accommodate the desired video/image data requirements for space applications. These approaches become increasingly important when data rates are limited by space loss (communications to Mars) or by a high demand for communications resources. One approach is to employ high speed data buffering so that image signals that are not "real-time" critical can be stored and retransmitted at a lower rate. It should be noted that one-way transmission times between Mars and Earth are somewhere between roughly 4 minutes (minimum distance) and 22 minutes (maximum distance). This, of course, alters the nature of two-way communications and also presents difficulties for controlling automated systems from Earth. For comparison, one-way (direct) transmission time between the moon and Earth is less than 2 seconds. Data storage devices can also be used to ease coverage and fade requirements for interplanetary missions. The number of planetary surface terminals and/or orbiting satellites could be reduced by recording mission data during outage periods and then transmitting the recorded data when communications is restored. Another approach for reducing transmission bit rate and data buffering requirements is to employ data compression techniques.

Lossy compression schemes can offer significant data rate reductions and yet still maintain an acceptable level of perceptual quality.

The technology assessment in [5] has identified uncompressed image and video data rate requirements to be on the order of 0.5-1000 Mbps, depending on the application. Image/video data can be classified into several categories as indicated in the following sections.

High Rate Video

As described in [5], it is desirable for many space missions to support high quality full motion video similar to standard National Television Systems Committee (NTSC) video. Supporting this level of video quality will require an uncompressed data rate of approximately 100 Mbps. As video standards evolve from the NTSC standard to High Definition Television (HDTV), it is desirable that future planetary missions have the capability to support the higher quality standard. Although, uncompressed data rates for HDTV are in the 1 GHz range, commercial interests are a driving force behind developing compression techniques for broadcasting high quality digital HDTV over the existing commercial television bandwidth allocation of 6 MHz.

Edited High Rate Video

A reduction in the video frame rate results in a lower quality video signal which in turn lowers the required uncompressed transmission rate to the 10-20 Mbps range. Remote monitoring and video teleconferencing are examples of applications where edited high rate video may be appropriate. As a result of the reduced frame rate associated with edited video, motion within an image is likely to result in objectionable distortions. Compression techniques can reduce the required transmission rate by a factor of 10 or more.

Low Rate Video

Certain video applications applications, such as remote monitoring, may have relatively low bit rate requirements. For example, an 8-bit, 512 by 512 monochrome signal transmitted at one frame per second requires an uncompressed transmission rate of slightly more than 2 Mbps and a compressed data rate of 0.2 Mbps or less.

Science Imaging Data

Due to the wide variety in technical requirements, it is difficult to quantify the image data requirements for scientific research. Estimates for uncompressed data rates vary widely from a low of 0.5 Mbps to as high as several gigabits per second [5]. It is unlikely that initial return links from Mars will be able to support raw data rates beyond that of compressed high rate video. Therefore, in many instances tradeoffs between resolution, frame rate and distortion levels (for compressed data) must be made. With scientific data, care must be exercised in choosing data compression schemes since the perceptual models used to achieve large compression ratios may not preserve the desired information content of scientific data. Since lunar and some space-to-space applications can support higher data rates, in these cases it may be desirable to transmit scientific data using lossless compression formats.

Multimedia Data

Multimedia involves the integration of text, high and low resolution computer graphics, audio and video into a coordinated information display. Advances in this area promise to provide substantial improvements in mission operations. For example, mission construction and maintenance procedures can be "interactively" displayed to crew members using

schematics, block diagrams and audio instructions. Furthermore, payload computers can be accessed and monitored by ground support facilities without flight crew intervention. A flight demonstration using Ku-band frequencies for transmitting 128 Kbps multimedia data to the space shuttle orbiter via commercial satellites has been proposed within the Tracking and Communications Division of NASA Johnson Space Center. Using compressed video and graphics it is possible to support multimedia with data rates as low as 128 Kbps. However, future multimedia applications, using higher resolution graphics and video, may require higher rates in the range of 0.5 to 5 Mbps.

Telerobotics Video

Telerobotics video relies on stereo images pairs to extrapolate the spatial information needed for autonomous or remote control. It is typically assumed that two high rate color video channels are needed for telerobotic video [5], resulting in an uncompressed transmission bit rate of roughly 200 Mbps. Some applications which require higher resolutions could push this figure into the Gbps range. Due to transmission delays high data rate telerobotics applications between Mars and Earth are impractical. Robotics applications will be forced to rely more heavily on autonomous control strategies thereby reducing the transmission bit rate requirements for telerobotic applications. Lunar communications links should be able to support the data rates required for advanced telerobotic applications. Furthermore, lunar transmission delays are sufficiently small so that lunar telerobotic applications are feasible.

Low Bit Rate Applications

Space missions must also support voice and command data applications. The data rate requirements for these uses are typically low when compared to the video/image case. Although water absorption in the atmosphere can lead to significant propagational losses at Ka-band, it appears that sufficient margin will exist to reliably maintain low bit rate applications for a Mars return link.

Link Margin Calculations

A link margin analysis is a valuable tool for determining a practical set of communications system parameters which are required to meet the system probability of error requirements. The link analysis consists of a "budget" which tabulates calculations regarding useful signal power and interfering noise power at the receiver. System gains and losses are totaled in terms of transmitter power, antenna gains, space loss, propagation losses, receiver noise figures, bit rate bandwidth and modulation/coding requirements. Link analyses of forward/return links for both a moon and a Mars link were performed using the proposed uplink frequency of 37 GHz and the proposed downlink frequency of 40 GHz. An example calculations for an interplanetary link is illustrated in Table 1. Several space-to-space scenarios operating near 26 GHz were also considered and are covered in the TCD report. The following sections give more detail regarding the calculations, models and assumptions used for the link margin results.

Transmitter Power

Size, weight and efficiency concerns give rise to transmitter power limitations for space based transmitters. This is especially true for space vehicles and satellites, however, planetary missions may have the payload capability to accommodate more powerful transmitters. Due to power requirements, the return link for interplanetary communications results in the

most stringent link requirements. Table 1 indicates a 75 Watt transmitter at a Mars ground station. The technology for such a transmitter appears to be feasible for near-term development. In [1, 6, 7] the authors point out that Hughes has developed a 60 GHz, coupled-cavity Traveling Wave Tube (TWT) amplifier which has dual mode operation at 30 and 75 Watts, and 40% efficiency. In general, manufacturing tolerances are tighter for higher frequencies and thus a 75 Watt transmitter at 40 GHz appears to be a practical assumption.

For lower power applications, such as some space-to-space applications, solid state amplifiers offer size and reliability advantages over TWTs. It appears that Monolithic Microwave Integrated Circuits (MMIC) amplifiers in the 0.1 to 5 Watt range will be technically feasible for Ka-band applications [8].

Antenna Gain

For most of the link margin calculations it is assumed that the transmitting and receiving antennas are parabolic-shaped reflector antennas each with efficiency $\eta = 0.55$ and having an antenna gain G given by

$$G = 10\log((\frac{\pi D}{\lambda})^2\eta) \text{ dBi} \quad ,$$

where D is the antenna diameter (in meters) and λ denotes the operating wavelength (in meters). However, for certain space-to-space applications we assume the use of a low gain antennas in order to avoid size and pointing difficulties. In these situations we assume that $G = 3$ dBi. One of the more ambitious antenna assumptions used in this study involves employing a 70 meter DSN antenna at 40 GHz. Achieving an antenna efficiency near $\eta = 0.55$ requires a root-mean-square (rms) surface deviation of approximately $\lambda/20$ m [9]. At 40 GHz this translates to a 375 μm rms surface deviation. Future work should investigate the technical specifications regarding the planned DSN upgrade to Ka-band frequencies (which was briefly mentioned in [4]). It is certainly possible that surface irregularities, combined with gravitational and wind forces, may reduce the antenna gain values used in our link calculations.

Space Loss

For space communications systems, space loss is the single largest loss in the system. The space loss, L_s, of a system is a loss in the sense that not all of the radiated energy from the transmitter is focused on the intended receiver. The calculation for space loss is given by

$$L_s = -20\log(\frac{4\pi r}{\lambda})(dB) \quad ,$$

where r denotes the path distance (in meters).

The worst case path loss for the Mars link is calculated based on a maximum separation distance of 2.675 Astronomical Unit, (AU) (1 AU = 1.496×10^5 km). As the Earth and Mars orbit the sun, the distance between the two planets varies from 0.374 to 2.675 AU. In all of the link margin calculations the worst case path loss is used. Therefore, at minimum separation there is over 16 dB excess margin. The average separation distance between Earth and Mars results in approximately 5 dB of available margin that can be used to increase link availability during rain conditions or to enable higher data rates during clear sky conditions. The worst case path loss for a lunar link is based on a maximum separation distance of 406,700 km.

Atmospheric Losses

The clear sky atmospheric losses due to water vapor and oxygen absorption are calculated using the Global Model [10]. Since the DSN sites operate at an elevation angle in the range of 10^o to 70^o, a conservative elevation angle of 20^o is employed in the propagation loss calculations. Note that the rain margins required for achieving low outage probabilities ($p < 1\%$) can be prohibitively large since fades in excess of 20 dB are not rare. However, for a $p = 5\%$ outage figure, the Global Model predicts a propagation loss of approximately 4 dB at a 20^o elevation angle (at Canberra, Australia). Of the three DSN sites, Canberra, Australia has the most stringent rain margin requirements and thus the overall network availability will be somewhat greater than 95%. Figure 4 illustrates bit-rate performance as a function of rain fade for several power levels.

The composition of oxygen and water vapor on Mars (by percent volume or number of molecules) is 10^{-1} and less than 10^{-1}, respectively. This is compared to Earth's composition of 21% and less than 1% for these gases [11]. Absorption losses through the Martian atmosphere due to oxygen and water vapor are therefore assumed to be insignificant at Ka-band frequencies. However, it should be noted that the Martian atmosphere contains relatively high levels of carbon dioxide, nitrogen and argon (up to 50% by volume). The absorption effects of these gases are also neglected in the propagation loss calculations.

Receiver Noise Temperature

For applications where the receiver antenna points toward cold space, the receiver noise temperature can have a significant effect on system performance. For example, under clear sky conditions the antenna temperature, T_a, associated with the DSN network is relatively low. Data in [10] indicates an antenna temperature of $T_a = 20$ K at 40 GHz and elevation angle $\theta = 90^o$. However, for lower elevation angles the increase in antenna temperature can have a significant effect on link performance. For example, [10] indicates an antenna temperature of $T_a = 120K$ at 40 GHz and $\theta = 10^o$. For low noise receivers (Noise Figures in the range of 0.25 – 1 dB), this increase in antenna temperature can result in a performance loss of 2 – 4.5 dB. Rain conditions also have a significant effect on antenna temperature. As mentioned previously, Figure 4 shows the channel bit rate performance as a function of rain fade. Since the Earth ground stations employ low noise receivers it is important to account for the increased antenna temperature when analyzing rain fade margins. Equations from [10] which relate the attenuation along an earth-space path and the antenna temperature are incorporated into the calculations used for Figures 4. As the path attenuation becomes large the antenna temperature asymptotically approaches the mean absorption temperature of the attenuating medium (in this case, 275 K).

For lunar links, there is a significant increase in antenna temperature due to solar energy reflected from the moon's surface. For operations near 40 GHz, the antenna temperature is calculated based on a model given in [10] and is calculated to be $T_a = 350K$. Since space-to-space links could have the Earth in the field of view, an antenna temperature of $T_a = 290K$ is employed for these applications.

Receiver noise figures are assumed to be in the range of 0.25 to 2.0 dB for Earth ground stations. Commercial space systems have ground receivers with noise figures in the range of 0.5 to 2.0 dB [12]. Miller [4] mentions that the 0.25 dB (17 K) noise figure is a planned upgrade to the DSN network. Space based receivers typically have larger noise figures. As a point of reference, at Ka-Band frequencies the Advanced Communications Technology

Satellite has a 3.4 dB noise figure, which is considered low by today's standards for a space based receiver. Assuming modest technology advances in space receiver designs, the noise figure for space receivers is assumed to be in the 2.0–4.0 dB range.

Modulation and Coding

The use of TWT amplifiers for providing the larger power levels required for Mars and lunar applications also has a bearing on the choice of modulation and multiple access schemes. In general, it is desirable to use a constant envelope modulation scheme (such as Multiple Phase Shift Keying (MPSK)) when a TWT amplifier is employed. Furthermore, Frequency Division Multiple Access (FDMA) schemes require TWTs to be "backed off" from their optimal power output levels in order to reduce intermodulation noise. Thus, when TWT amplifiers are used, Time Division Multiple Access (TDMA) schemes are more desirable for providing multiple access.

The performance of a digital communications system can be improved by the use of coding. Coding involves the incorporation of redundancy (usually in the form of parity bits) in an effort to detect and correct errors that are induced by a noisy channel. Concatenated coding schemes are often used in space applications since they offer relatively high coding gains and also use interleaving to combat burst errors. A concatenated Reed-Solomon (RS) and convolution code (with Viterbi decoding) can provide coding gains in the 8.5–9.5 dB range for a bit error probability of $P_e = 10^{-8}$ [13]. Thus the required $\frac{E_b}{N_0}$ for BPSK is reduced from 15.2 dB (without coding) to 5.7 dB (with coding). The use of certain data compression schemes requires more stringent bit error probability requirements and hence we use $P_e = 10^{-8}$ instead of the often used $P_e = 10^{-5}$ value. However, the several of the space-to-space applications do not assume a concatenated scheme, but rather assume a convolution code and $P_e = 10^{-5}$. Since this approach requires $\frac{E_b}{N_0} = 9.6$ dB, and the coding gain is roughly 5 dB, the difference between the two approaches does not greatly effect the results of the link margin calculations. Of course, a concatenated scheme provides improved P_e performance at the expense of encoding/decoding complexity. Advances in integrated circuit technologies continue to push the data rates that can be achieved by various concatenated coding schemes to higher levels. We have assumed that coding gains of 9.5 dB can be achieved for the lunar data rates in the near future. Presently, a concatenated RS and short block code can achieve coding gains of up to 7.5 dB and can be implemented for high data rates [13]. A more detailed description of various concatenated coding schemes can be found in [14].

Note that traditional coding schemes achieve coding gain via bandwidth expansion. For example, a concatenated (255,223) RS code and a $\frac{1}{2}$ rate, constraint length 7 code mentioned in [1] requires that over 50% of the bits transmitted over the channel to be parity bits. In 1982, Gottfried Ungerboeck published a technique called Trellis-Coded Modulation (TCM) [15] that combines multilevel modulation and coding to achieve coding gain without the usual bandwidth expansion. As an extension of the summer's work, we are investigating the use of a concatenated RS and TCM scheme which achieves coding gain with only a modest degree of expansion, all of which is due to the RS code. Further work will continue in this area as we investigate the level of coding gain which can be achieved using this technique.

Pointing Loss

Pointing loss is a potentially major issue for implementing an interplanetary communi-

cations system. For this study we use a "placeholder" value of 0.5 dB to represent pointing loss. Using an approximation given in [16] we have that the 3 dB beamwidth, θ (which is in degrees) is given by

$$\theta = \frac{75\lambda}{D} \quad .$$

At 40 GHz, $\theta = 0.008^0$ for a 70 m DSN antenna, and $\theta = 0.056^0$ for a 10 m MST antenna. Thus, the beam spot size at the MST is approximately $5d_m$, where d_m denotes the diameter of Mars. The beam spot size at the DSN is roughly $18d_e$, where d_e denotes the diameter of the Earth. A lunar link with a 10 m Earth GS and a 2 m LST will result in beam spot sizes of $1/10d_m$ and $1/10d_e$, respectively, where d_m denotes the diameter of the moon. Further work in this area is needed to more accurately quantify pointing losses for interplanetary applications.

Polarization Loss

In this study a placeholder value of 1 dB is used to represent polarization loss. The underlying assumption is that both the transmitting and receiving antennas employ circular polarization. The 1 dB placeholder value is based on assumptions used by [1] and [4].

Implementation Loss

Given the lack of hardware development at Ka-band frequencies, it is not possible to precisely determine all of the link parameters. For this reason a placeholder value of 2 dB is used to represent the expected losses due to the practical implementation of each communications system. Note also that the link calculations in this report are designed to achieve a 3 dB circuit margin.

Discussion and Conclusions

As mentioned previously, Tables 1 provides a "typical" link margin calculation for an interplanetary communications link. Note that the Earth to Mars link requires significantly more power due to the fact that space based receivers have significantly higher noise temperatures. Due to the difference in space loss, a lunar link can support significantly higher data rates when compared to the Mars application.

Figures 3 and 4 illustrate various tradeoffs between channel bit rate, transmitter power, rain fade margins and transmitter/receiver antenna size. Example nominal link parameter values and circuit margin equations are given in Table 1. Unless otherwise stated within the figure or the figure caption, the graphs shown in Figures 3 and 4 are based on the default parameters found in the appropriate tables mentioned previously. Both of the figures are calculated assuming a 3 dB circuit margin. Figures 3 illustrates a trade between transmitter antenna diameter and transmitter power for the Mars return link. Figure 4 illustrates the degrading effect on system performance that can be expected due to rain fades. It is desirable for the interplanetary communications systems to have the capability to accommodate several data rates. During fading conditions communications can be maintained by switching to a lower data rate. Furthermore, the space loss variability for a Mars mission is significant and can be exploited to provide rain margin and/or increased data rates.

Table 1: Mars to DSN Link Margin Calculations (37 GHz)

xmt power, dBw	18.8		75	Watts
circuit loss, dB	-2.0	estimate		
antenna gain, dBi	63.1	dish size	6	m
		frequency	37	GHz
		efficiency	0.55	
EIRP, dBw	79.9			
			2.675	AU
space loss, dB	-295.9	range	400180000	km
		freqency	37	GHz
polarization loss, dB	-1.0	estimate		
rcv antenna gain, dBi	86.1	dish size	70	m
		efficiency	0.55	
propagation loss, dB	-1.4	Global Model		
pointing loss, dB	-.5	estimate		
Total Received Power (TRP), dBW	-132.8			
System Noise Temperature	19.2	Antenna Temp	32	K
		Noise Figure	0.5	dB
		rcv circuit loss	0.1	dB
Receive G/T, dB/K	66.9			
Boltzman's constant, dBW/K/Hz	-228.6		1.38E-23	W/K/Hz
noise spectral density	-209.4	System Noise Temp+Boltzmann's constant		
TRP/NSD, dB	76.6			
bit rate bandwidth, dBHz	66.0		4	Mbps
implementation loss, dB	-2.0	estimate		
modulation loss, dB	0			
coding gain	9.5	concatenated RS-convolutional		
Received Eb/N0	18.1	TRP/NSD-bandwidth+imp. loss+mod loss+coding gain		
Required Eb/N0	15.2	BPSK, BER=10e-8		
Link Margin	2.9	Received Eb/N0-Required Eb/N0		

Table 1: Mars to DSN Link Margin Calculations (37 GHz)

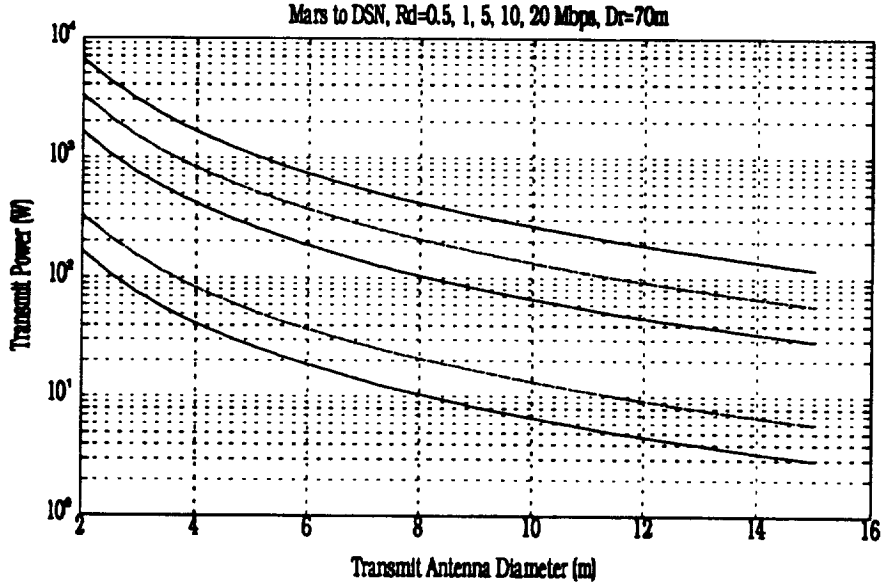


Figure 3: Mars transmitter power (W) versus transmit antenna diameter for channel bit-rates of 0.5, 1, 5, 10 and 20 Mbps and a 70 m receive antenna. Link parameters not specifically listed above are shown in Table 1. The curves are in descending order by bit-rate with the 20 Mbps curve on top and the 0.5 Mbps curve on the bottom.

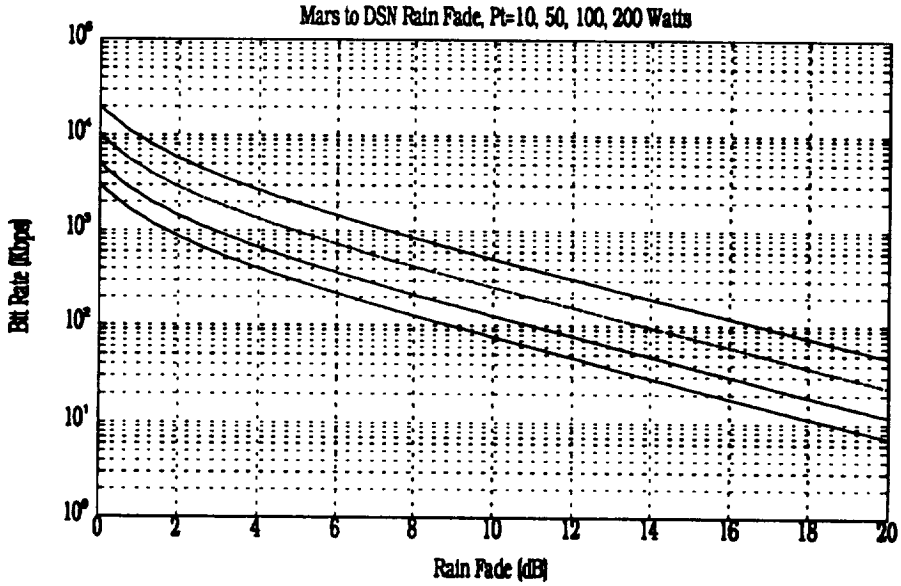


Figure 4: Bit-Rate (Kpbs) as a function of rain fade (dB), for Mars transmitter power values 10, 50, 100 and 200 watts. Link parameters not specifically listed above are shown in Table 1. The curves are in ascending order by power level with the 10 watt curve on bottom and the 200 watt curve on top.

References

[1] Denise S. Ponchak, John E. Zuzek, Rodney L. Spence, and Wayne A. Whyte, Jr. "Millimeter wavelength communications applications for the SEI". In *AIAA/NASA/OAI Conference on Advanced SEI Technologies*, Cleveland, OH, September 4–6 1991. Paper Number 91-3589.

[2] Rodney L. Spence (NASA–Lewis Research Center). "A comparison of optical technologies for a high data rate Mars link". In *AIAA/NASA/OAI Conference on Advanced SEI Technologies*, Cleveland, OH, September 4–6 1991. Paper Number 91-3469.

[3] D. S. Ponchak and J. E. Zuzek and W. A. Whyte, Jr. and R. L. Spence and P. Y. Sohn. A technology asseessment of alternative communications systems for the Space Exploration Initiative. In *AAIA Space Programs and Technology Conference*, Huntsville, Alabama, September 25–27 1990. NASA TM-103243.

[4] Edward F. Miller (NASA-Lewis Research Center). "Overview of Ka-band communications technology requirements for the space exploration initiative ". In *AIAA/NASA/OAI Conference on Advanced SEI Technologies*, Cleveland, OH, September 4–6 1991. Paper Number 91-3424.

[5] Wayne A. Whyte, Jr. and Khalid Sayood. "Data compression for full motion video transmission ". In *AIAA/NASA/OAI Conference on Advanced SEI Technologies*, Cleveland, OH, September 4–6 1991. Paper Number 91-3533.

[6] James A. Dayton, Jr. (with the NASA-Lewis Research Center). "Status of Ka-band TWT transmitter technology". In *AIAA/NASA/OAI Conference on Advanced SEI Technologies*, Cleveland, OH, September 4–6 1991. Paper Number 91-3621.

[7] James A. Dayton, Jr. (with the NASA-Lewis Research Center). "High power highe efficiency millimeter wavelength traveling wave tubes for high rate communications from deep space". In *AIAA/NASA/OAI Conference on Advanced SEI Technologies*, Cleveland, OH, September 4–6 1991. Paper Number 91-3593.

[8] Edward J. Haugland. "high efficiency Ka-band MMIC amplifiers. In *AIAA/NASA/OAI Conference on Advanced SEI Technologies*, Cleveland, OH, September 4–6 1991. Paper Number 91-3425.

[9] John D. Kraus. *Antennas.* McGraw-Hill, New York, second edition, 1988.

[10] Louis J. Ippolito, R. D. Kaul, and R. G. Wallace. *Propagation Effects Handbook for Satellite Systems Design: A Summary of Propagation Impairments on 10 to 100 GHz Satellite Links With Techniques for System Design.* NASA Reference Publication 1092(03), third edition, 1983.

[11] Richard Goody and James Walker. *Atmospheres.* Prentice-Hall, Englewood Cliffs, New Jersey, 1972.

[12] Bernard Sklar. *Digital Communications Fundamentals and Applications.* Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

[13] Leon W. Couch II. *Digital and Analog Communication Systems.* MacMillan, New York, fourth edition, 1993.

[14] Shu Lin and Daniel J. Costello, Jr. . *Error Control Coding: Fundamentals and Applications.* Prentice-Hall, Englewood Cliffs, New Jersey, 1983.

[15] Gottfried Ungerboeck. "Channel coding with multilevel/phase signals. *IEEE Transactions on Information Theory*, IT-28:55–67, January 1982.

[16] Timothy Pratt and Charles W. Bostian. *Satellite Communications*. John Wiley and Sons, New York, 1986.

LOADING, ELECTROMYOGRAPH, AND MOTION DURING EXERCISE

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Beth A. Todd, Ph.D.
Academic Rank:	Assistant Professor
University & Department:	The University of Alabama Department of Engineering Science and Mechanics
NASA/JSC	
Directorate:	Space and Life Sciences
Division:	Medical Sciences
Branch:	Space Biomedical Research Institute
JSC Colleague:	Linda C. Shackelford, M.D.
Date Submitted:	August 11, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

A bicycle ergometer system has been developed to determine forces acting in specific muscles and muscle groups for both cycling and isometric exercise. The bicycle has been instrumented with encoders, accelerometers, and load cells. A harnessing system has been developed to keep subjects in place during isometric exercise. EMG data will also be collected with electrodes attached to various muscles on the subject's leg. Data has been collected for static loading and will be collected for cycling in both an earth-based laboratory and on the KC-135. Once the data is analyzed, the forces will be entered into finite element models of bones of the lower extremities.

A finite element model of the tibia-fibula has been generated from the experimental subject's MRI data. The linear elastic isoparametric brick elements representing the bones are connected by linear elastic isoparametric shell elements placed at the locations of ligaments. Models will be generated for the calcaneus and the femur. Material properties for the various tissues will be taken from the literature. The experimentally determined muscle forces will be applied to the models to determine the stress distribution which is created in the bones.

INTRODUCTION

Under weightless conditions, astronauts have exhibited a tendency to develop osteoporosis in their lower extremities. For locomotion, they push off the walls of the orbiter with their arms, but the usual forces of a gravitational environment are no longer applied to their legs. With the projected loss of bone in the femoral neck, missions over 11 months in duration will put 40-year-old astronauts in a moderate fracture risk category equivalent to a 70-year-old. According to Wolf's Law, bone adapts to functional demands by remodeling to reflect the distribution of effective stresses. [1] Exercise countermeasures which impose an earth-like stress distribution on the bones of the lower extremities will have a positive affect on bone density loss.

This report describes the initial phase of a long-term project in which various exercise devices will be studied to determine their effectiveness in the prevention of bone density loss. A bicycle ergometer has been instrumented with accelerometers, position encoders, and load cells. From this instrumentation and electrodes attached to a subject's leg, force data will be taken both on earth and in the weightless environment of the KC-135. This data will be analyzed to determine the forces acting in each of the major muscle groups of the lower extremities. This dynamic force data will then be entered into finite element models of the tibia-fibula, femur, and calcaneus to calculate the stress distribution throughout these bones in both the one-g and zero-g environments.

EXPERIMENTAL EQUIPMENT

The data acquisition system and the equations for analyzing the accelerations and forces on the bicycle are described in report #11 by Dr. Figueroa. Traditionally, this type of data is collected by photographing an activity using two cameras mounted at a sufficient distance to view the entire scene. The data acquisition system was developed for the bicycle ergometer with the confined quarters of space vehicles in mind.

ELECTROMYOGRAPHY

To determine forces acting on individual muscles, electromyographic data is used. During the motion of bicycling, the primary movements are ankle extension and flexion. However, numerous other muscles in the leg are also activated. These secondary muscles, such as the peroneus longus, contribute to such things as stabilization of the ankle joint. Countering forces are applied to the bones which are not easily measured as external work. The primary hypothesis of this project is that secondary muscles apply loads to the bone which can be

beneficial in preventing bone loss.

Data for determining the EMG/force relation for a variety of calf muscles was collected. With the assistance of Gary Dudley, Ph.D. and William Norfleet, M.D., measurements were made on individual muscles with corresponding torques measured on the LIDO.

EMG electrodes were placed on a female subject's left leg at the following locations: peroneus longus, lateral gastrocnemius, medial gastrocnemius, soleus, flexor hallux, peroneus brevis, extensor digitorum longus, and tibialis anterior. Measurements were made on the left calf with the knee moving from extension to flexion at rates of 30 and 90 degrees per second while the subject was in a seated position. Measurements were repeated for the left ankle with the subject in a kneeling position, and they were also made for unilateral and bilateral stances and toe raises with eyes open and closed. These measurements were repeated after local lidocaine injection about the motor point of each muscle group. After the first injection, all toe extensors and a majority of the tibialis anterior were eliminated. The second injection eliminated the peroneus longus; after this injection, the subject had no dorsi-flexion. The third injection eliminated the soleus and left minimal function remaining in the calf. The toe flexors remained intact. Injection four may have eliminated the tibialis posterior. More will be known about the effect of that injection when the data is analyzed. The fifth and final injection was in the adductor canal proximal to the peroneal nerve branch. All motor functions below the knee except the supination-plantar flexion (tibialis posterior, flexor digitorum longus, and flexor hallux longus) were eliminated. Complete use of the calf muscles returned to the subject a few hours after the last injection.

A sample plot of the raw EMG data is shown in Figure 1. The square curve is the signal associated with the recorder being activated. The signal following that is the EMG.

EMG and force data were also collected off of the left leg of the same subject under static conditions. A harnessing system was developed to keep the subject connected to the bicycle ergometer while the subject exerted maximal effort with the instrumented leg.

Additional data will be collected from this subject while cycling. This data will be collected both in an earth-based laboratory and on the KC-135.

DATA ANALYSIS

The experimental data will be analyzed with the computer package MATLAB. The EMG data will be analyzed using traditional procedures. [2] The

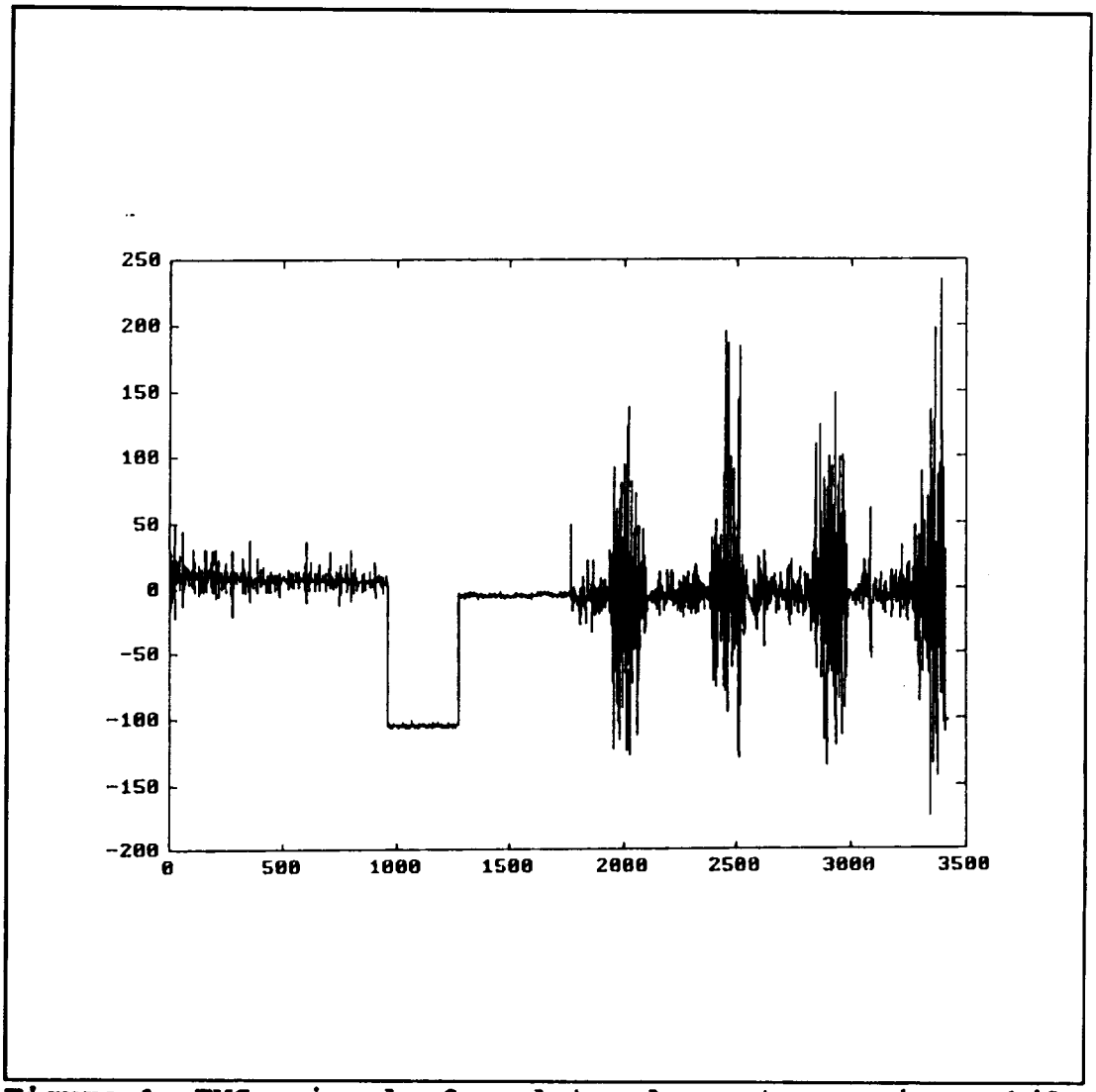


Figure 1 EMG signal for lateral gastrocnemius while subject is free-wheeling on ergometer.

accelerometer, load cell, and encoder data will be analyzed with the equations developed in report #11. [3] Forces acting at specific locations of the bones of the lower extremities in different situations will be calculated from this analysis. These resulting forces will be used in the finite element model to calculate the stress distributions in the calcaneus, femur, and tibia-fibula.

FINITE ELEMENT MODELING

The bones of greatest interest are the tibia-fibula, femur, and calcaneus. To determine the stress distribution in each of these anatomical structures, finite element models are needed for each of them.

Geometry

The geometry for the finite element models is based on data taken with Magnetic Resonance Imaging (MRI). Additional details around the joints of the bones were determined from x-rays. The use of these imaging techniques will allow the generation of models which will correspond to individual experimental subjects as opposed to a general-purpose generic model.

Eighty-two images of the subject's right leg were made. Slices 9 to 38 were 256x256 bit transverse images of the calf spaced 12 mm apart. Slices 7, 8, and 51 to 82 were 256x256 bit transverse images of the thigh spaced 12 mm apart. The other images were sagittal and coronal views of the body. Data from the images were transferred into *.bitmap (*.bmp) files.

PAINTBRUSH was found to be the most useful IBM-compatible software for determining nodal coordinates for the final element model. The *.bmp files were read into PAINTBRUSH, and the bone was identified. Twelve nodal points were identified around both the inner and outer edges of the bone for each slice. Using the Zoom feature with Cursor Position turned on, pixel coordinates for these points were recorded. One of the MRI slices for the calf is shown in Figure 2.

The pixel coordinates for each of the nodes were converted into mm with MATLAB. (Conversion factors are 0.488 for the calf and 0.781 for the thigh.) Then an input file was constructed for the PC-based finite element program PRIME. [4] The geometry for six 36-mm long regions was generated with PRIME. The geometry had to be generated in these regions due to the large amount of memory required by the whole model as compared to the 4 Mb of RAM which were available on the IBM-compatible 386 used for this project. The complete model will be generated on a DECStation 5000/240 with 40 Mb of RAM running at approximately 50 mips and solved on that computer as well as an IBM RS6000.

The model of the tibia-fibula is nearing completion. Currently the model contains 1242 nodes and 696 elements. The bones are constructed from 8-node linear orthotropic isoparametric brick elements. They are connected with 4-node and 3-node linear isotropic isoparametric shell elements which are placed at the locations of the ligaments. [5] Shell thicknesses vary from 5 to 8 mm depending on their location. [6,7]

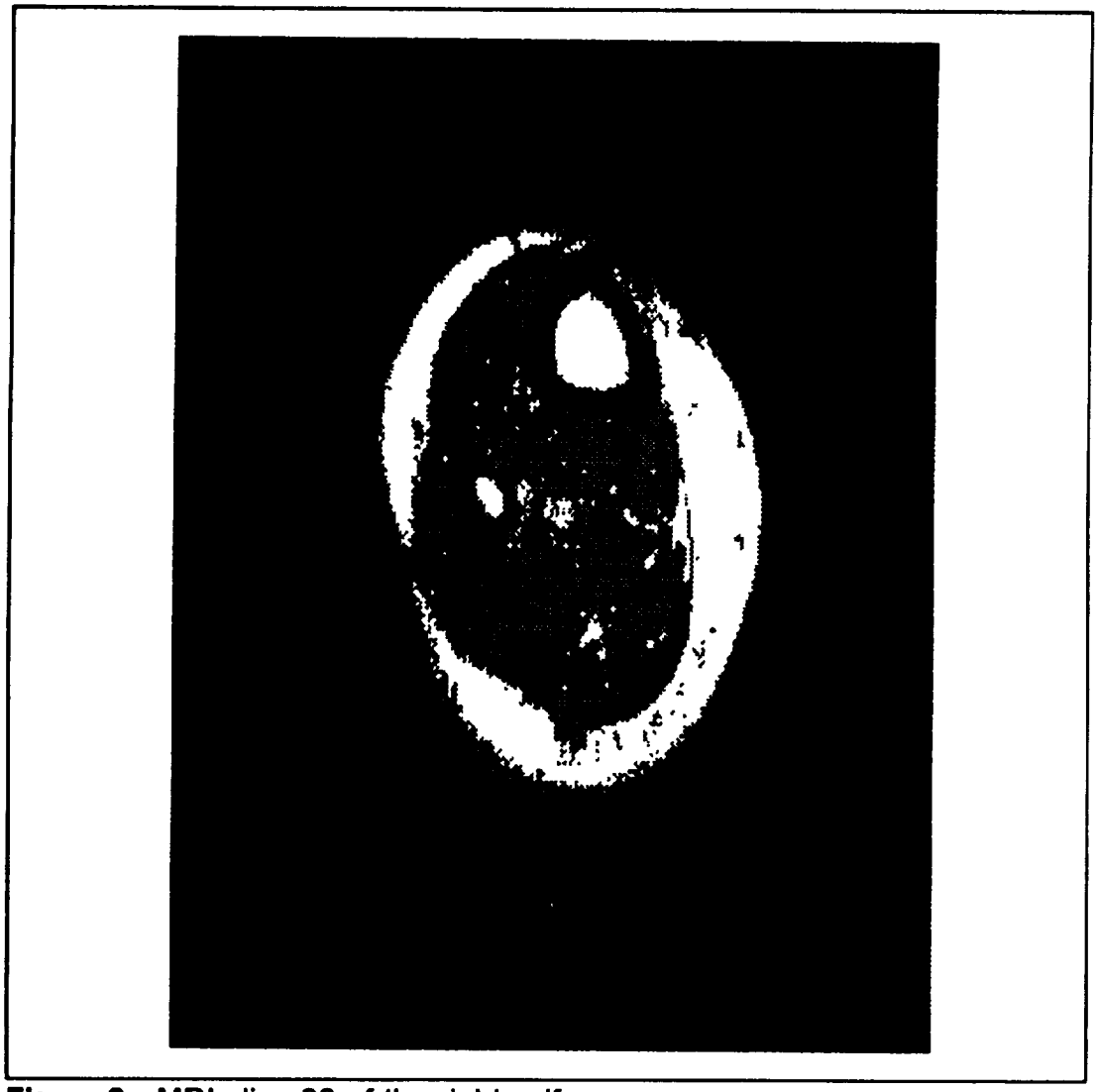


Figure 2 MRI slice 30 of the right calf.

Material Properties

Material properties for the models will be taken from the literature. For the elements representing ligaments, Young's Modulus, E, was set to 8.33 MPa with Poisson's Ratio, ν, as 0.49 based on published data. [8]

For bone, there are numerous articles available which discuss their material properties. [9-29] Values vary as a function of storage method, testing conditions, specimen configuration, specimen orientation, anatomical origin, porosity, density, and ash content. [20] Some values appear to be influenced by the aims of the research. [30] Some published values appear to conflict with other published values. [30] Some values [12,13] fall outside of theoretical limits, such as

Poisson's Ratio calculated to be greater than 0.5. [31,32] Clearly, the literature must be understood thoroughly so that accurate material properties can be chosen. Currently, E = 15500 MPa and ν = 0.28. [33] It is expected that a more thorough examination of the literature will result in modifications to these values.

In particular, the orthotropic properties of bone will be considered. Reilly and Burstein have determined five elastic constants which describe the orthotropic properties of bone: E_1, E_3, G_{13}, ν_{13}, and ν_{12}. [8,13] Due to the aims of this research, work defining material properties as a function of density will be studied. [20] Upon placing portions of femoral shaft in a four-point bending test, Keller et al. found an exponential relationship between Young's modulus and apparent density. However, the value of the exponent varies with the anatomical location of the specimen. Also, at a future time, some smaller portions of the bones may be modeled to include some of the microstructural information which is becoming better understood. [34]

Loading

Forces will be applied to the bones at the locations of muscle attachment. The directions of the experimentally determined forces will be same as those which are actually applied to the bones.

Boundary Conditions

Other researchers have modeled experiments of bones of the lower extremities, and their boundary conditions have matched the boundaries of their experimental apparatus. In these models of the lower extremity bones, points away from the loads will be constrained in all degrees of freedom. Results will be examined in light of fictitious reactions around these constrained points. The models may be reanalyzed with different boundary conditions to eliminate these fictitious reactions from the results.

CONCLUSIONS AND RECOMMENDATIONS

A bicycle ergometer system has been developed to determine forces acting in specific muscles and muscle groups for both cycling and isometric exercise. Data has been collected for static loading and will be collected for cycling in both an earth-based laboratory and on the KC-135. Once the data is analyzed, the forces will be entered into finite element models of bones of the lower extremities.

A finite element model of the tibia-fibula has been generated from the experimental subject's MRI data. Models will be generated for the calcaneus and the femur. Further work needs to be done to automate the process for generating geometry from MRI slices.

REFERENCES

1. Bentzen SM, Hvid I, Jorgensen J: Mechanical strength of tibial trabecular bone evaluated by x-ray computed tomography. *Journal of Biomechanics* 20:743-752, 1987.

2. Winter DA: *Biomechanics and Motor Control of Human Movement* 2nd Ed.,Wiley-Interscience, 1990.

3. Figueroa JF: Loading, electromyograph, and motion during exercise. Final Report, NASA/ASEE Summer Faculty Fellowship Program, 1993.

4. Berry KJ: *PRIMEGENIX™ User Manual* Geometrics, Inc., Flint, MI, 1993.

5. Williams PL and Warwick R: *Gray's Anatomy* 36th Ed., W.B. Saunders Company, 1980.

6. Trent PS, Walker PS and Wolf B: Ligament length patterns, strength, and rotational axes of the knee joint. *Clinical Orthopaedics and Related Research* 117:263-270, 1976.

7. Netter FH: *The CIBA Collection of Medical Illustrations--Volume 8. Musculoskeletal System--Part 1. Anatomy Physiology, and Metabolic Disorders*, pp. 94-101.

8. Woo SL-Y, Orlando CA, Sato S, Camp JF, Gomez MA and Akeson WH: The effects of aging, temperature and post mortem storage on ligament tensile behavior. *1984 Advances in Bioengineering*, ed.: RL Spilker, ASME, 1984.

9. Fung YC: *Biomechanics--Mechanical Properties of Living Tissues*, Springer-Verlag, 1981.

10. Vannah WM and Childress DS: An investigation of the three-dimensional mechanical response of bulk muscular tissue--experimental methods and results. *ASME Symposium on Computational Methods in Biomechanics*, 1988.

11. Ashman RB: Chapter 5--Experimental techniques. *Bone Mechanics*, ed. SC Cowin, CRC Press, 1989.

12. Handbook of Engineering in Medicine and Biology, eds. DG Fleming and BN Feinberg, CRC Press, pg. 254, 1976.

13. Reilly DT and Burstein AH: The elastic and ultimate properties of compact bone tissue. *Journal of Biomechanics*, 8:393-405, 1975.

14. McElhaney JH, Fogle JL, Melvin JW, Haynes RR, Roberts VL, and Alem NM: Mechanical properties of cranial bone. *Journal of Biomechanics*, 3:495-511, 1970.

15. Cowin SC, Van Buskirk WC and Ashman RB: Chapter 2--Properties of Bone. *Handbook of Bioengineering*, eds. R Skalak and S Chien, McGraw-Hill, pp. 2.1-2.27, 1987.

16. Reilly DT, Burstein AH and Frankel VH: The elastic modulus for bone. *Journal of Biomechanics*, 7:271-275, 1974.

17. Martin RB and Ishida J: The relative effects of porosity, density, mineralization and fiber orientation on the tensile properties of cortical bone. *Orthopaedic Research Society*, 33rd Annual Meeting, 1987.

18. Mizrahi J, Condax P, Edwards WT and Hayes WC: Finite element analysis of the parameters affecting strength reduction of vertebral bodies. *Orthopaedic Research Society*, 37th Annual Meeting, 1991.

19. Odgaard A and Linde F: The underestimation of Young's modulus in compressive testing of cancellous bone specimens. *Journal of Biomechanics*, 8:691-698, 1991.

20. Keller TS, Mao Z and Spengler DM: Young's modulus, bending strength, and tissue physical properties of human compact bone. *Journal of Orthopaedic Research*, 8:592-603, 1990.

21. Martin RB: Determinants of the mechanical properties of bones. *Journal of Biomechanics*, 24-1:79-88, 1991.

22. Hollister SJ, Goldstein SA, Kuhn JL, and Matthews LS: Correlation of finite element stress predictions with trabecular bone remodeling *in vivo*--effects of mesh refinement and mechanical property assumptions. *Orthopaedic Research Society*, 34th Annual Meeting, 1988.

23. Rao AA and Dumas GA: Influence of material properties on the mechanical behaviour of the L_5-S_1 intervertebral disc in compression--a nonlinear finite element study. *Journal of Biomedical Engineering*, 1991.

24. Faulkner KG, Cann CE, and Hasegawa BH: Effect of bone distribution on vertebral strength--assessment with patient-specific nonlinear finite element analysis. *Radiology*, 179:669-674, 1991.

25. Keyak JH and Skinner HB: Three-dimensional finite element modelling of bone--effects of element size. *Journal of Biomedical Engineering*, 14:483-489,

1992.

26. Odgaard A, Hvid I and Linde F: Compressive axial strain distributions in cancellous bone specimens. *Journal of Biomechanics*, 22:829-835, 1989.

27. Marom SA and Linden MJ: Computer aided stress analysis of long bones utilizing computed tomography. *Journal of Biomechanics*, 23:399-404, 1990.

28. Turner CH, Cowin SC, Rho JY, Ashman RB and Rice JC: The fabric dependence of the orthotropic elastic constants of cancellous bone. *Journal of Biomechanics*, 23:549-561, 1990.

29. Keyak JH, Meagher JM, Skinner HB, and Mote CD: Automated three-dimensional finite element modelling of bone--a new method. *Journal of Biomedical Engineering*, 12:389-397, 1990.

30. Natali AN and Meroi EA: A review of the biomechanical properties of bone as a material. *Journal of Biomedical Engineering*, 11:266-276, 1989.

31. Gere JM and Timoshenko SP: *Mechanics of Materials*, 3rd ed., PWS-Kent, 1990.

32. Juvinal RC, *Engineering Considerations of Stress, Strain, and Strength*, McGraw-Hill, pp. 53-56, 1967.

33. Todd BA: *Finite Element Analysis of the Stress Distribution within the Seated and Supine Human Buttocks, In Vivo*, Ph.D. Dissertation, University of Virginia, 1992.

34. Goldstein SA, Feldkamp LA, Ku JL, Champlain FW, Zand M, Jesion G, Ciarelli M and Matthews LS: Architectural effects on trabecular bone material properties. *Orthopaedic Research Society*, 32nd Annual Meeting, 1986.

N94-25381

MODELING OF THE GROUND-TO-SSFMB LINK NETWORKING FEATURES USING SPW

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	John C. Watson, Ph.D.
Academic Rank:	Instructor
University & Department:	Amarillo College Electronics Technology Dept. Amarillo, Texas 79178
NASA/JSC	
Directorate:	Engineering
Division:	Tracking & Communications
Branch:	Systems Engineering
JSC Colleague:	Laura Hood
Date Submitted:	August 3, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

This report describes the modeling and simulation of the networking features of the Ground-to-Space Station Freedom Manned Base (SSFMB) Link using COMDISCO Signal Processing WorkSystem (SPW). The networking features modeled include the implementation of Consultative Committee for Space Data Systems (CCSDS) protocols in the multiplexing of digitized audio and core data into Virtual Channel Data Units (VCDUs) in the Control Center Complex and the demultiplexing of VCDUs in the onboard Baseband Signal Processor. The emphasis of this work has been placed on techniques for modeling the CCSDS networking features using SPW. The objectives for developing the SPW models are to test the suitability of SPW for modeling networking features and to develop SPW simulation models of the Control Center Complex and Space Station Baseband Signal Processor for use in end-to-end testing of the Ground-to-SSFMB S-band Single Access Forward (SSAF) Link.

INTRODUCTION

The Ground-to-SSFMB Link Networking Features Simulation Model shown in Figure 1 includes the segments of the SSAF link [1] in which networking features of the CCSDS protocol are implemented. Audio data received within the Control Center Complex are formatted into Bitstream Protocol Data Units (BPDUs). Data Management System (DMS) Core Data are received through CCSDS variable-length packets which are multiplexed into a fixed-length Multiplexing Protocol Data Unit (MPDU). In the absence of sufficient core data to complete a MPDU data zone, a fill data CCSDS packet is appended. The Virtual Channel Data Unit (VCDU) Frame Multiplexer receives the BPDUs and MPDUs for insertion into the VCDU data zone which accommodates one BPDU or MPDU. Synchronous audio data receive priority transfer over core data. In the complete SSAF link the VCDUs would be encrypted with the Data Encryption Standard (DES) and Reed-Solomon encoded to form Coded Virtual Channel Data Units (CVCDUs). The CVCDUs would then be scrambled with a pseudorandom noise (PN) sequence to form Channel Access Data Units (CADUs) for transmission to Space Station Freedom. However, since this study primarily concerns the networking features of the SSAF link, the link will be observed immediately after the assembly of VCDUs. At this point, VCDU bit errors are randomly introduced to observe their impact on the proper reconstruction of recovered CCSDS packets and audio data within the Space Station Baseband Signal Processor. The LNA, Transponder, and portions of the Space Station Baseband Signal Processor which perform Frame Sync/Random Sequence Demodulation, Reed-Solomon Decoding, and DES decryption will be bypassed in this analysis which concerns primarily the networking features of the CCSDS protocol as applied to Space Station Freedom.

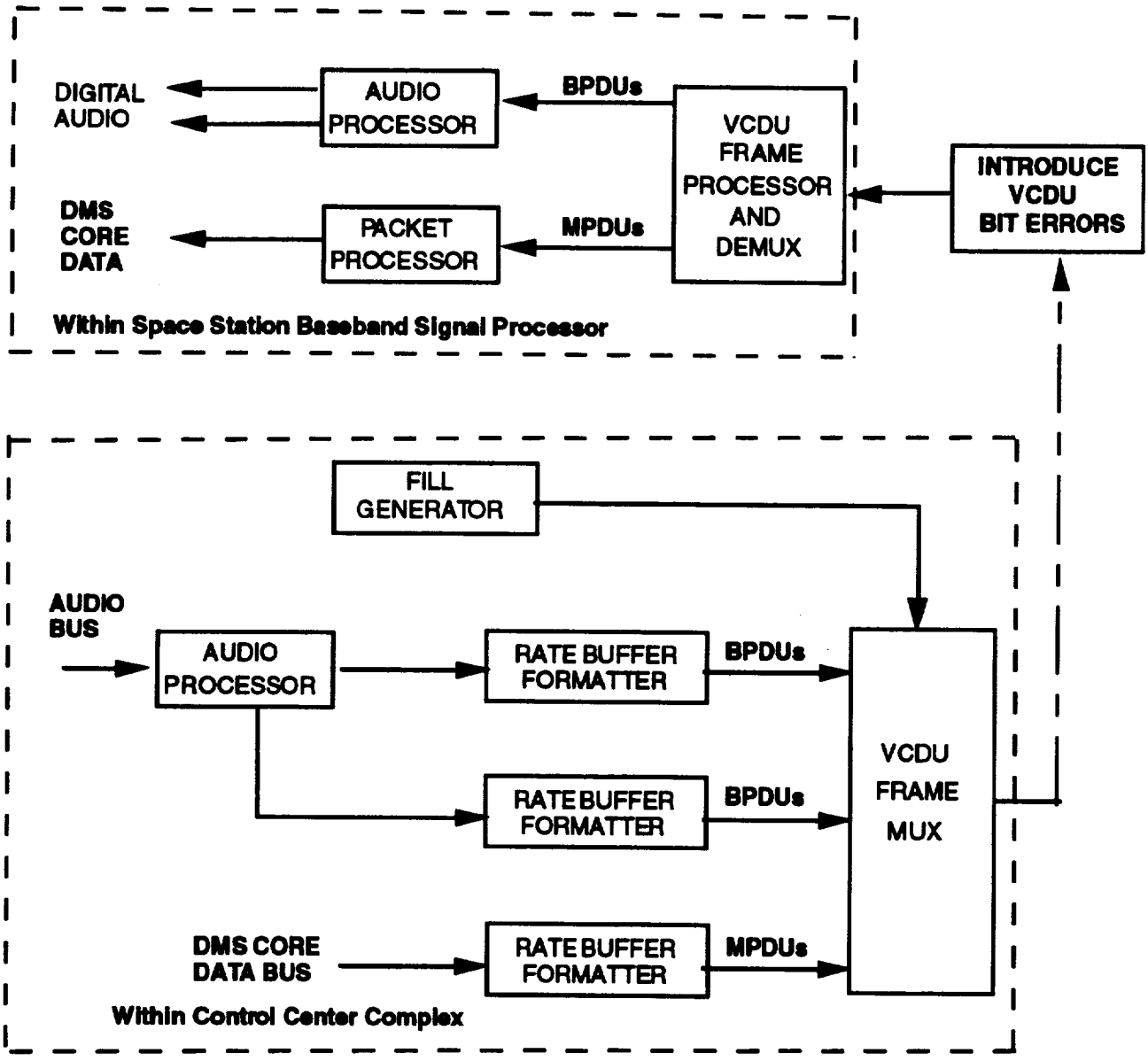


Figure 1. Ground-to-SSFMB Link Networking Features Simulation Model

The VCDUs are received by the VCDU Frame Processor and Demultiplexer where the BPDUs and MPDUs are recovered. The Audio Processor recovers the digital audio data channels from the BPDUs. The Packet Processor recovers the CCSDS packets by reading the CCSDS packet headers to reassemble the variable-length CCSDS packets sent from the Control Center Complex. The Ground-to-SSFMB Link Networking Features Simulation Model shown in Figure 1 was modeled using SPW [2]. SPW is a block-oriented simulator for communication systems. The complete SPW simulation model for the multiplexing and demultiplexing of VCDUs has been described in [3].

SPW Modeling of Networking Features

A primary objective for developing the Ground-to-SSFMB Link Networking Features simulation model was to determine the suitability of COMDISCO Signal Processing WorkSystem (SPW) as a network simulation tool for packet telemetry systems. Major advantages in using SPW are its signal processing capabilities whereby signals can be analyzed and the multi-rate feature which allows processes to operate at different rates during simulation time. For example, the processing rate for a data packet generator can be adjusted so that the packet can be generated within the maximum allowable frame time. When using the multi-rate feature of SPW, care must be taken in generating the required timing signals used to coordinate the various processes used in the simulation model. It can be especially confusing when shifting data between vector and serial forms. SPW appears to have some basic software limitations, however, which must be overcome to develop networking simulation models. Each of these basic limitations will be discussed with suggested techniques for working with the limitation.

SPW Vectors Must Be Fixed-Length

SPW processes data in serial or vector form. SPW vectors used to model transfer frames, protocol data units, and data packets must be fixed-length. The fixed-length vector constraint becomes a problem when attempting to generate the CCSDS variable-length packets which are inserted into the MPDU data zone. For this reason, the CCSDS packet length must be fixed for each CCSDS packet generator modeled in SPW.

Parameters Must Be Deterministic

The parameters used to define SPW block instances cannot be random. For example, in describing a digital data source, the duration of the data transmission cannot be specified as a random variable. This constraint poses a problem in the modeling of random data packet duration and interarrival time. To overcome the deterministic parameter constraint, special modules must be created to provide the necessary statistics for data generation. The Digitized Audio Data Packet Generator is one such module in which a Poisson Impulse Generator, Poisson Random Number Generator, Simple Counter, Comparator, and R-S flip-flop were used to generate digital audio data packets with random data set duration and packet interarrival time.

SPW Library Lacks Basic Functional Modules

Many of the fundamental building blocks for digital systems are omitted in the SPW Library. For example, SPW includes no modules for performing decimal-to-binary conversion and binary-to-decimal conversion. Converters were developed to provide for the generation of CCSDS packet headers and the demultiplexing of CCSDS packet data. SPW does not include shift registers for

sorting through serial data. The Circular Buffer included in the SPW library has a fixed-length vector and does not provide a reset for clearing the buffer. Some of the basic modules for processing vectors were not included in SPW. Modules for performing XOR, AND, OR, and vector switching were exported from COMDISCO to develop the Ground-to-SSFMB link networking SPW simulation model.

Conversion Between SPW Vector and Serial Data Can Be Misleading

Packet telemetry system simulation requires the use of vectors to model transfer frames, protocol data units, and packets. The vectors can be generated by performing a 16-bit scalar to vector conversion using a Scalar Join Vector Block or loading serial data into a Circular Buffer. If a Scalar Join Vector Block is used, the serial-to-vector conversion is accomplished in one simulation iteration. When using a Circular Buffer the length of the Circular Buffer and the data rate of the source loading the buffer must be considered in determining the number of iterations required to convert the serial data to vector data. The timing signal used to drive the Vector Sink Hold pin which allows vector data to be viewed after the completion of the simulation must be synchronized correctly before the vector signal will be displayed. Conversion from serial-to-vector and from vector-to-serial data can introduce phase shifts which make data comparison tricky. For these reasons, the packet telemetry system simulation model uses a vector data representation of the digitized audio and DMS core data. By keeping the Control Center Complex and Baseband Signal Processor digitized audio and core data signals as vector signals, a data comparison can be implemented in the SPW simulation model without compensating for throughput delays.

Ground-to-SSFMB Link Networking Features SPW Model

A SPW simulation model was developed to simulate the networking features of the Ground-to-SSFMB link. The detailed description of the simulation model is described in [3]. The SPW model for simulating the Ground-to-SSFMB Link Networking Features shown in Figure 2 consists of three high-level modules. The SSAF Control Center Complex (ssafccc) module simulates the part of the Control Center Complex which generates VCDUs from digitized audio and DMS core data sources for transmission to Space Station Freedom. The Noise module introduces random bit errors into the VCDU frames. The probability of channel bit error can be adjusted within the Noise module. The SSAF Baseband Signal Processor (ssafbsp) module receives the VCDUs for demultiplexing into audio and DMS core data. The high level modules are linked through the following signals:

vcdu	Virtual Channel Data Unit vector
dms	Data Management System core data vector

audio Digitized audio data vector

ccsdshdr CCSDS header vector

mpduhdr Multiplexing Protocol Data Unit header vector

These vector signals are required to provide an error check of frame, header, and data errors introduced by the Noise source.

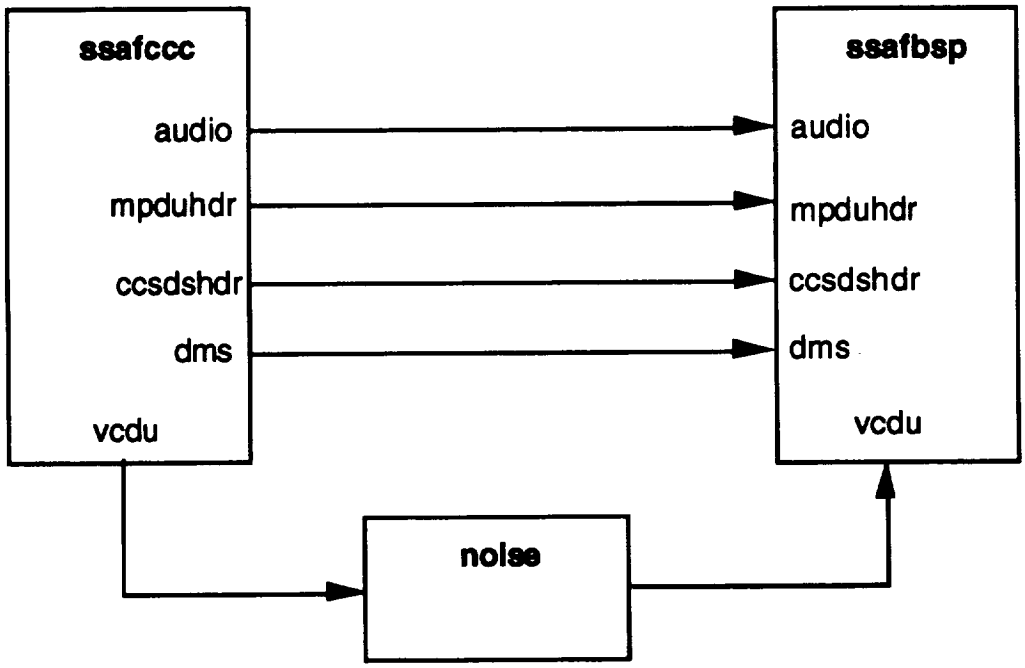


Figure 2. Ground-to-SSFMB Link Networking Features SPW Model

SSAF Control Center Complex SPW Module

The SSAF Control Center Complex SPW module details are shown in Figure 3. Although SPW includes signal generation modules, these modules cannot be set with random variable parameters. The AUDIO module was developed to create a digital data packet generator which has random packet interarrival times and packet length durations. The AUDIO module generates digitized audio data packets which have exponential packet interarrival times and Poisson-distributed packet durations. The BPDUGEN module receives the digitized audio data packets from the AUDIO module and formats them into BPDUs. The DATAFLAG signal in Figure 3 provides a logical true to indicate when audio data are present.

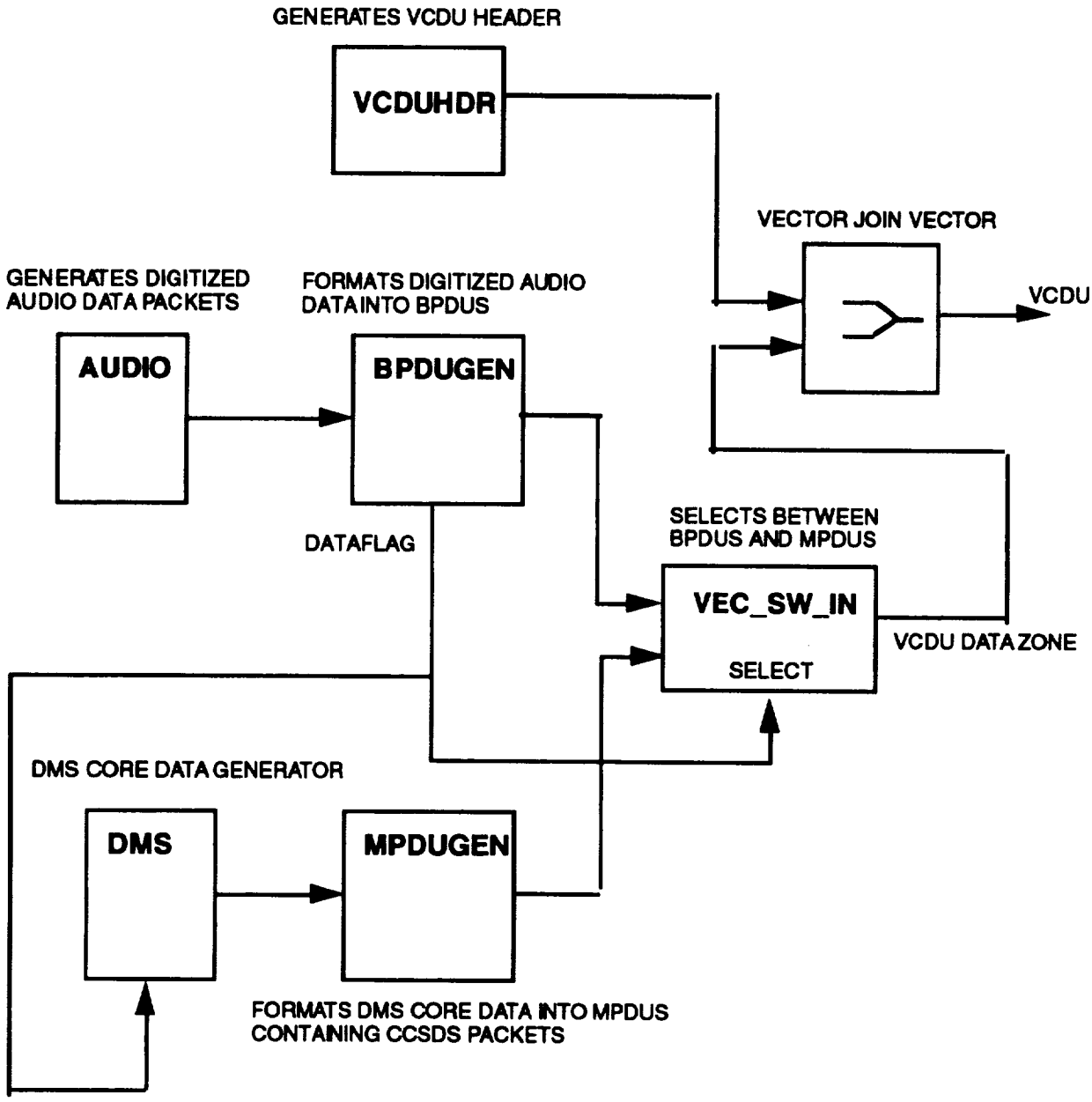


Figure 3. SSAF Control Center Complex Networking Features SPW Module

The DATAFLAG signal drives the BPDU/MPDU Vector Select Switch. BPDUs receive priority transfer over MPDUs since they contain synchronous data. The DMS module generates the Data Management System core data vector to be transferred through the MPDU Generator in the form of CCSDS packets. Conversion of the DMS core data into a vector format provides a data comparison with the recovered DMS core data in the onboard Baseband Signal Processor. The MPDUGEN module inserts the DMS core data into CCSDS

packets. The CCSDS packets are then inserted into the MPDU data zone. Ideally, the MPDUGEN module would generate MPDUs containing variable-length CCSDS packets for transferring the DMS core data. However, the SPW fixed-length vector constraint for SPW block instances complicates the generation of variable-length CCSDS packets. The MPDU Generator module generates MPDUs containing four CCSDS packets. Each CCSDS packet has a 360-bit data field and 48-bit CCSDS packet header for a total CCSDS packet length of 408 bits. Four CCSDS packets complete the 1632-bit MPDU data zone. The MPDUGEN module encodes the Application Process ID, Packet Sequence Number, and Packet Length into CCSDS packet header information. The MPDU and CCSDS headers are provided through output ports to provide header error detection in the onboard Baseband Signal Processor (ssafbsp) module. The Vector Switch-In (VEC_SW_IN) module was exported from COMDISCO to provide vector selection. This module was constructed by COMDISCO and was not part of the SPW library. The Vector Switch-In module selects between a BPDU and MPDU for insertion in the VCDU Data Zone. The VCDUHDR module generates the 48-bit VCDU header and 64-bit VCDU Insert Zone. The VCDU header and Insert Zone were analyzed together in the SPW module since they reside in every SSAF VCDU. This was done to simplify the model although in future models the Crypto Sync code within the VCDU Insert Zone may require analysis separate from the VCDU header. The VCDU header, VCDU Insert Zone, and VCDU Data Zone are combined by the Vector Join Vector block to form the Control Center Complex VCDU.

Noise Module

The Noise Module shown in Figure 4 inserts bit errors into the VCDU Frame by inverting a randomly selected bit. The Random Data Generator in Figure 4 drives the Select signal on the Vector Switch Out block. The Vector Switch Out block either sends the VCDU directly to the Baseband Signal Processor or routes the VCDU through the Extract Component block which removes a randomly selected bit for inversion. The Probability of Zero parameter within the Random Data Generator sets the probability of VCDU frame error. The Noise Generator block in Figure 4 models a Uniform Distribution with limits 0 through 1759 for selecting the index number of the VCDU frame bit to be clobbered. The Insert Component block places the inverted frame bit back into the VCDU.

SSAF Baseband Signal Processor SPW Module

The SSAF Baseband Signal Processor SPW Module shown in Figure 5 models the networking features of the onboard Baseband Signal Processor. The VCDUs from the Control Center Complex are received through the VCDU port of the module. The Vector Split block separates the VCDU Header from the received VCDU. The recovered VCDU Header is compared with a mask of the Control Center Complex VCDU Header. The Vector Compare block, developed by Larry Johnson of Lockheed Engineering and Science Company, provides a logical

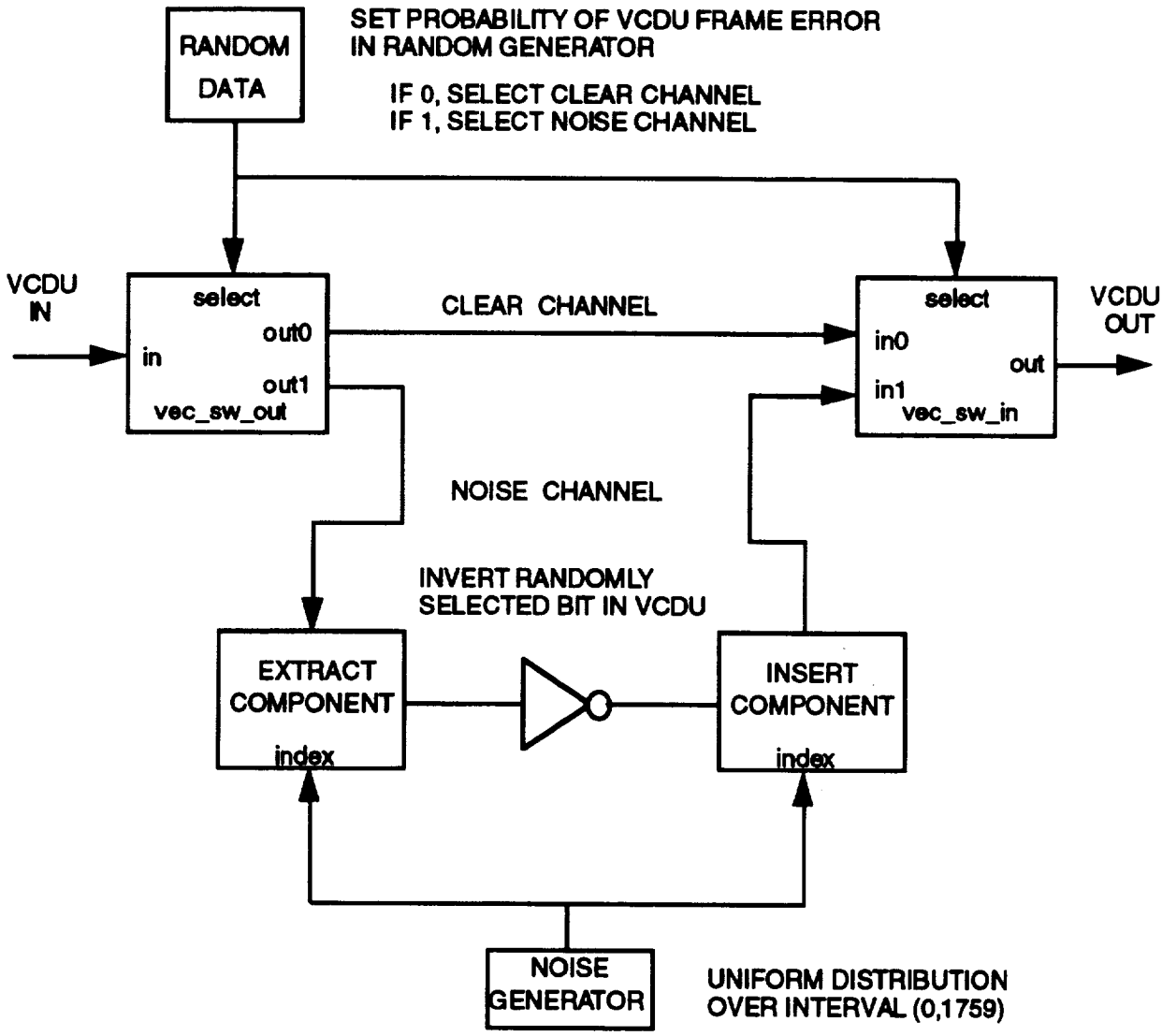


Figure 4. Noise Module

"true" signal output when two input vectors are the same. The signal output of the VCDU Header Vector Compare block (#1) is inverted to provide a VCDU Header error signal. The 1648-bit VCDU Data Zone is removed from the received VCDU at the same time the VCDU Header is removed. The VCDU Data Zone will contain either a BPDU containing digitized audio data or a MPDU containing CCSDS data packets. The VCDU Data Zone is simultaneously tested for a BPDU header and a MPDU header as shown in Figure 5. The BPDU header Vector Compare block (#2) compares the first 16 bits of the VCDU Data Zone with the BPDU Header Mask. The BPDU header has a fixed bit pattern (hexadecimal 3FFF). If the BPDU header Vector Comparator determines that a valid BPDU header is present, then the signal from the Digitized Audio Vector comparator (#3) will be valid. The Digitized Audio Vector comparator tests the

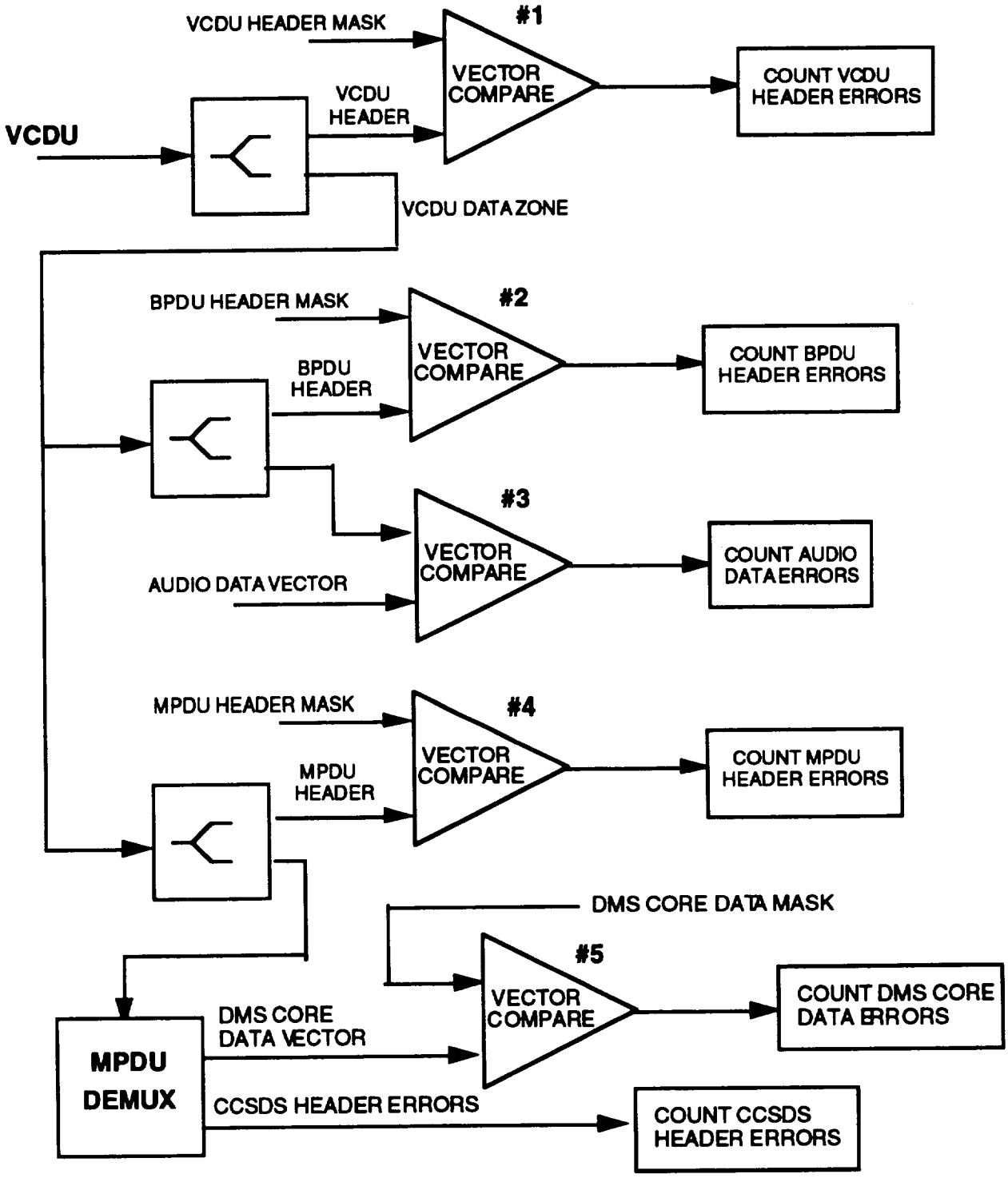


Figure 5. SSAF Baseband Signal Processor SPW Module

BPDU Data Zone with the Control Center Complex digitized audio data vector signal. The BPDU header error signals and Digitized Audio data error signals are recorded using the same procedure used to record VCDU header errors. The MPDU header Vector Compare block (#4) determines whether a valid MPDU header is present in the VCDU Data Zone by comparing the first 16 bits of the VCDU Data Zone with the Control Center Complex MPDU Header vector. If a valid MPDU header is detected, the MPDU Data Zone will be demultiplexed by the MPDU Demultiplexer block. The MPDU Demultiplexer block generates the CCSDS packet header error signal and recovers the DMS core data vector from the MPDU Data Zone. The DMS core data Vector Compare block (#5) compares the recovered DMS core data vector with the Control Center Complex DMS core data vector to determine if bit errors were introduced into the core data. DMS core data errors and CCSDS Header errors are recorded during the simulation. Not shown in Figure 5 are the logical gates required to ensure that a received BPDU will not be recorded as a MPDU header/data error and vice versa.

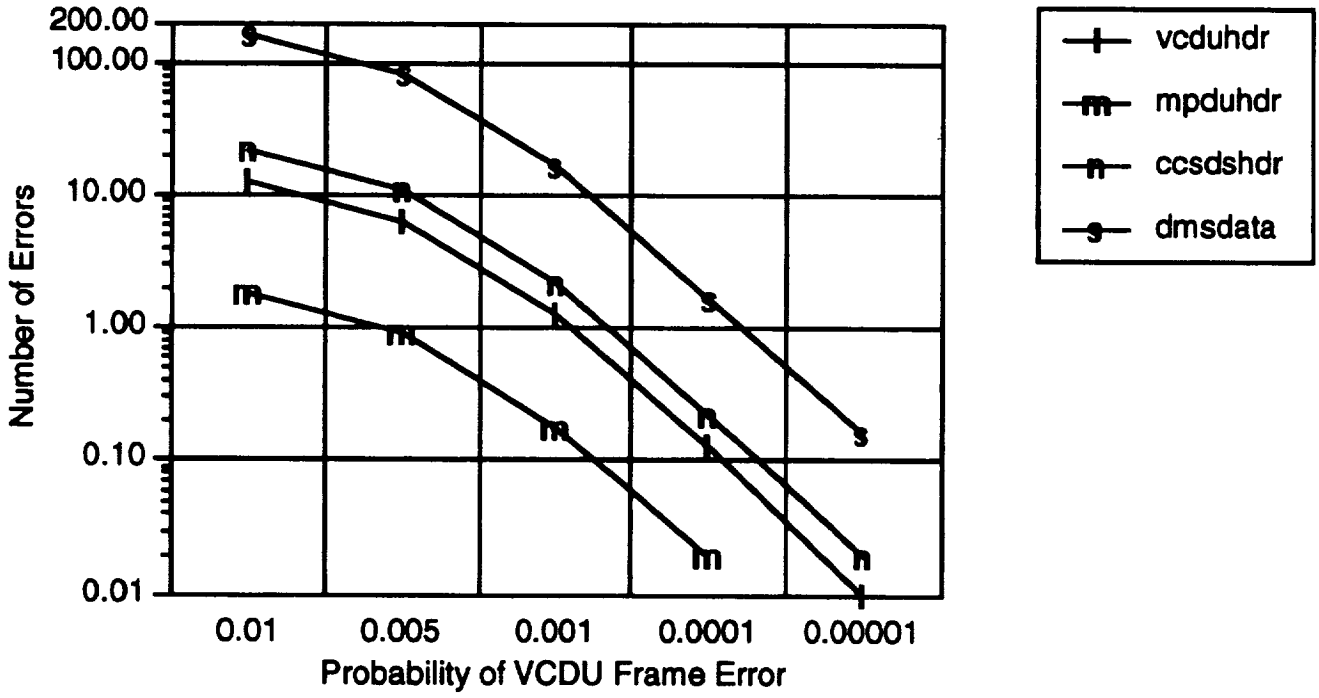


Figure 6. Theoretical Error Performance for MPDU's for 20,000 VCDU Frames

Theoretical Performance

An objective for developing the SPW simulation model for the Ground-to-SSFMB Link was to determine the impact of noise on the recovery of digital audio and core data from the demultiplexed VCDUs in the Baseband Signal Processor. The probability that a header receives a bit error within a VCDU frame increases with

the number of bits in the header. The Theoretical Performance plot shown in Figure 6 displays the expected number of header and data errors over the observation of 20,000 VCDU frames for different VCDU frame error probabilities. The performance plot was developed assuming that one bit error will occur within the 1760-bit VCDU if the VCDU has a frame error. The performance plot in Figure 6 was made for the recovery of MPDU's where each MPDU contains four CCSDS packets transporting DMS core data. A performance plot for BPDU recovery would be similar to Figure 6 except that the number of expected audio data bit errors would differ. The theoretical performance for BPDU and MPDU generation would require defining the percentage of audio data transmission relative to DMS core data transmission.

Simulation Performance

The Ground-to-SSFMB Link SPW Model was simulated for 20,000 VCDU frame generations at different VCDU frame error probabilities to produce the SPW Simulation Performance Plot shown in Figure 7. The Ground-to-SSFMB SPW Simulation Model generates BPDU's whenever audio data packets are received. To observe the generation of MPDU's, the Digitized Audio data generator packet interarrival time within the SPW model was set so that DMS core data transport would dominate to provide a comparison with the Theoretical Performance plot given in Figure 6 for MPDU's. The SPW simulation model records an integer number of errors for the VCDU header, MPDU header, CCSDS packet header,

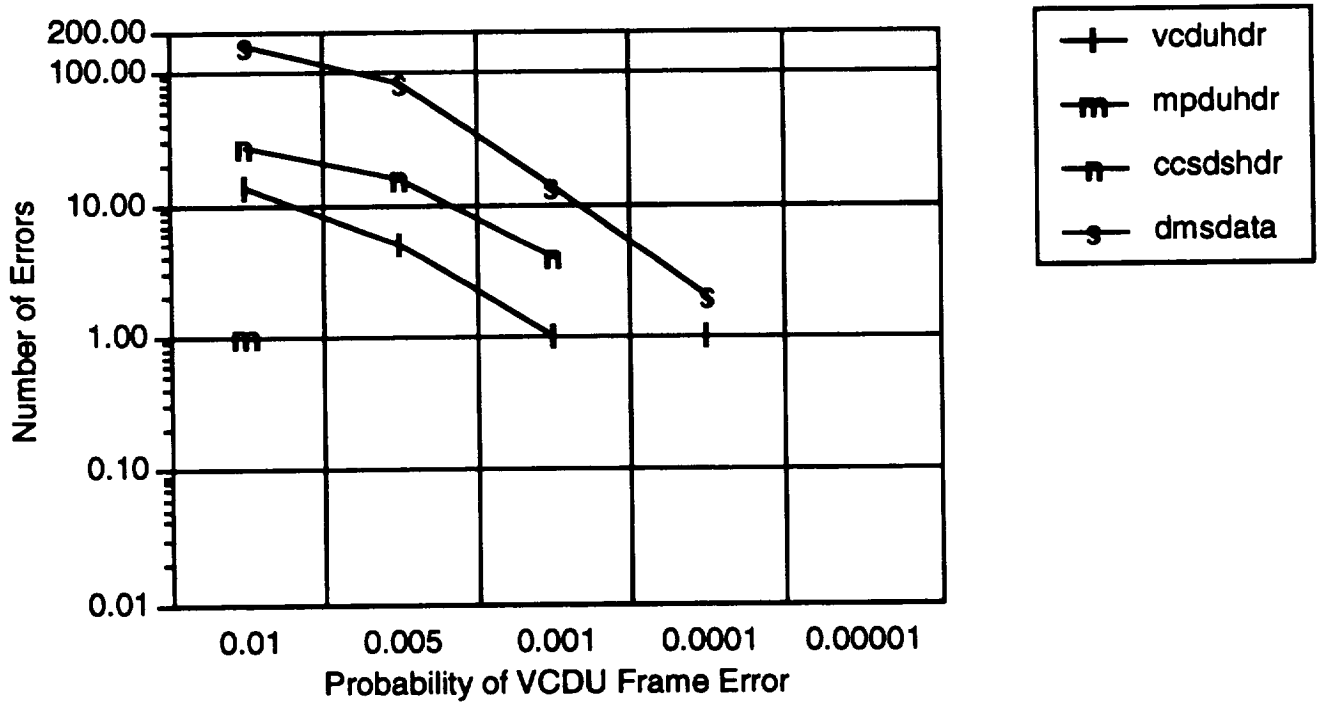


Figure 7. SPW Simulation Error Performance for MPDU's for 20,000 VCDU Frames

and DMS core data error. SPW simulation results agree with the theoretical performance results for an integer number of errors. Longer simulation durations would enable the performance plots to be presented with integer error counts providing a more accurate comparison with the theoretical performance plot for lower VCDU frame error probabilities.

VCDU Frame Simulation Using Delayed Unit Pulse Signals

An alternative method for simulating the networking features of the CCSDS-based packet telemetry system suggested by Karim Alhussiny of Lockheed Engineering and Science Company is to represent packet header and data as unit pulses delayed in time with respect to their appearance in the VCDU. In the delayed unit pulse signal VCDU frame simulation model, each block of header and data information is represented as a unit pulse whose duration is equal in time to the number of bits in the block. For example, in Figure 8, the 48-bit VCDU header has been modeled as a 48-sample unit pulse. The 64-bit VCDU Insert Zone has been modeled as a 64-sample unit pulse delayed 48 samples so that it occurs at the proper time with respect to the VCDU frame. The 16-bit BPDU header has been modeled as a 16-sample unit pulse delayed 112 samples to synchronize the occurence of the BPDU header with the VCDU frame. The unit pulse representation of the VCDU frame blocks models the case in which all bits of each block are set to one. This representation does not reflect variations in VCDU header information, but does provide a method for noise analysis which simulates rapidly. Regardless of how the header bits are initially set, a bit inversion due to noise will have the same impact in terms of performance in the

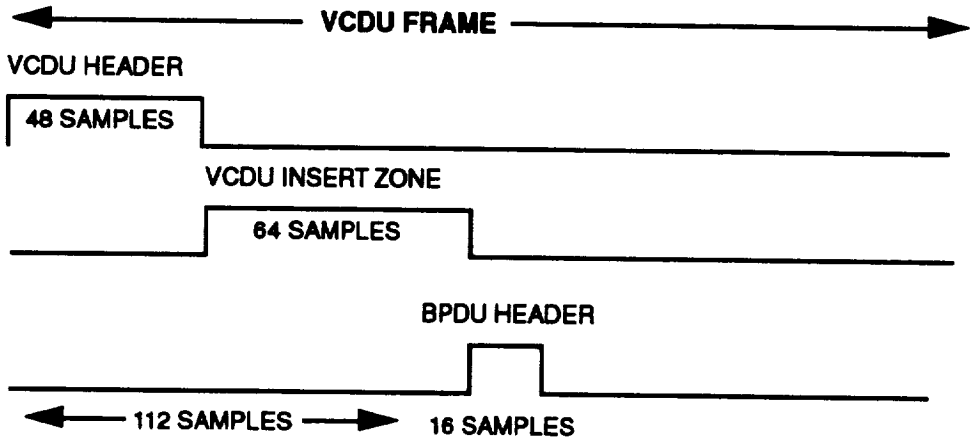


Figure 8. Unit Pulse Representation of VCDU Frame Blocks

presence of noise assuming the header bits do not include a coding scheme. The complete VCDU frame can be modeled as the set of delayed unit pulses ORed together. This generates a unity signal over the 1760 samples contained in the SSAF VCDU. The generated pulses are used to activate HOLD signals on the comparator blocks so that a bit error will not be interpreted over the wrong time

interval for the VCDU block under observation. Ten unit pulse generators are required to generate the VCDU header and Insert Zone (112-bits), MPDU header (16 bits), CCSDS packet headers (48 bits), and CCSDS packet data zones (360 bits). The serial data model was developed to provide a comparison with the vector version of the SPW model for the Ground-to-SSFMB link networking features. The SPW simulation model which decribes the VCDU frame in terms of delayed unit pulses executes very rapidly. The simulation of 20,000 VCDU frames required less than one minute to simulate compared to almost one hour simulation time required for the vector SPW model described earlier in this report. The performance results shown in Figure 9 for the SPW Delayed Unit Pulse Model compare favorably with the theoretical results for VCDU frame errors shown in Figure 6. Although the unit pulse representation of the VCDU frame does not accurately describe header details for signal analysis, it does provide a test for estimating the effect of noise on clobbered VCDU frames without requiring time-intensive simulations. As in the vector SPW model, the Delayed Unit Pulse Simulation model does not represent variable-length CCSDS packets.

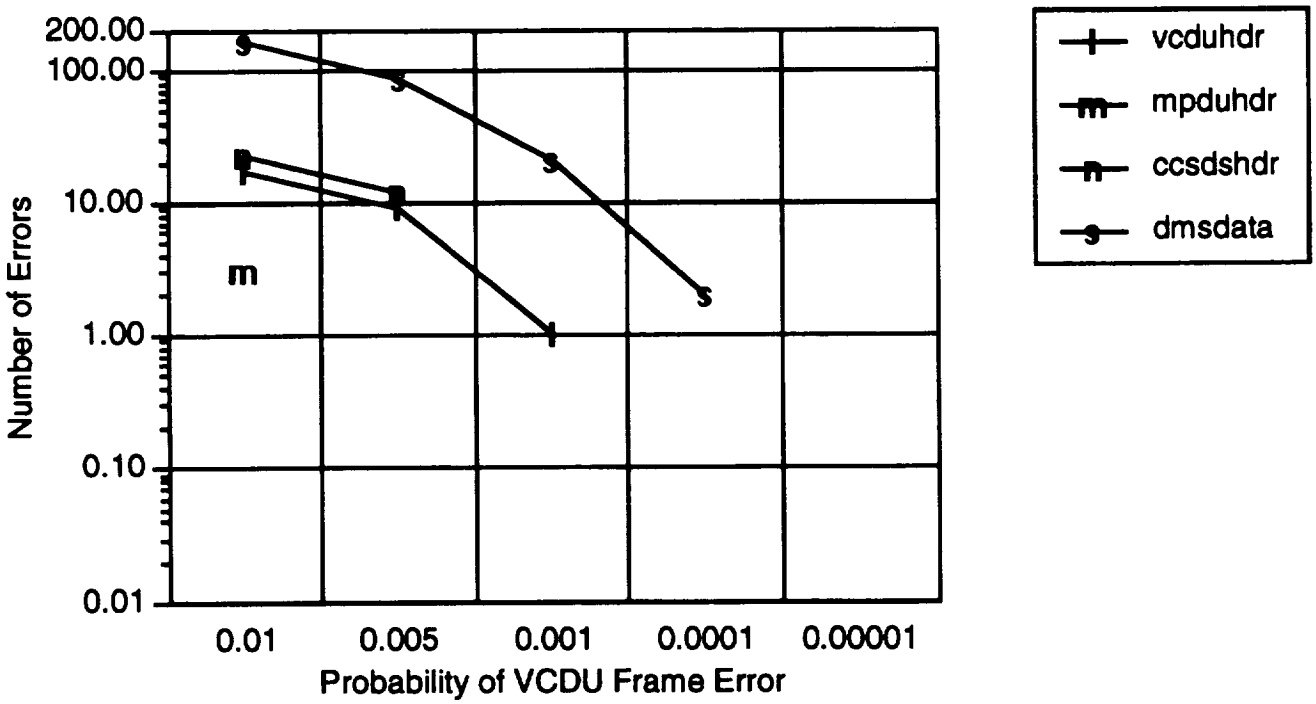


Figure 9. SPW Delayed Unit Pulse Simulation Model Results for 20,000 VCDU Frames

Conclusion

The primary SPW software constraint which must be overcome is the ability to randomly vary the vector lengths so that the generation of variable-length CCSDS data packets can be simulated. The model described in this work assumed the CCSDS data packet generation scenario in which the maximum number of CCSDS packets would be inserted within the MPDU. The effects of channel noise will be paramount whenever the condition of greatest frame overhead occurs, which will be when the CCSDS packets have minimum length. Although the simulation of the SPW model compared well with theoretical performance results, the fixed-length CCSDS packet specification was not the desired design requirement for the SPW model. Ideally, the SPW model would read the CCSDS packet header information within the Baseband Signal Processor to reassemble the segmented data residing in the individual CCSDS packets. Although the SPW model encodes the desired packet header information in binary format, it does not use this information to demultiplex the data once the MPDU has been extracted from the VCDU data zone. The packet header information was not used since the vector length for the data field could not be specified as a variable which could be adjusted by the binary-to-decimal conversion of the packet length field within each CCSDS data packet header. If the SPW fixed-length vector constraint could be circumvented, perhaps through developing specially coded vector block instances, SPW simulation models could be developed which more accurately model the networking features of a CCSDS-based packet telemetry system.

REFERENCES

1. Space Station Freedom Program Interface Control Document SPACE STATION FREEDOM (MANNED BASE) TO GROUND, Lockheed Engineering & Sciences Company, November 1992.

2. COMDISCO Signal Processing WorkSystem (SPW), Version 3.0, 1992.

3. Ground-to-SSFMB Networking Features SPW Simulation Model, John C. Watson, 1993.

N94-25382

SOFTWARE ENGINEERING METHODOLOGIES AND TOOLS

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Lawrance (Larry) M. Wilcox
Academic Rank	Assistant Professor
University & Department	Southeastern Oklahoma State University (SOSU) Computer Science/Information Systems Department Durant, OK 74701
NASA/JSC	
Directorate:	Engineering
Division:	Tracking and Communications
Branch:	Systems Engineering
JSC Colleague:	Sally Stokes, EE
Date Submitted:	July 30, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

The Engineering Profession is reported to be the second oldest profession of humankind. The first engineering discipline was mechanical engineering to support the building of war machines. Over the years many engineering disciplines have developed including chemical, electronic, etc. Common to all engineering disciplines is the use of rigor, models, metrics and predefined methodologies.

Recently a new engineering discipline has appeared on the scene, called Software Engineering. For over thirty years computer software has been developed and the track record has not been good. Software development projects often miss schedules, often are over budget, often does not give the user what is wanted and it has defects. One estimate is there are 1 to 3 defects per 1000 lines of deployed code. More and more systems are requiring larger and more complex software for support. As this requirement grows the software development problems grow exponentially. It is believed that software quality can be improved by applying engineering principles.

Another compelling reason to bring the engineering disciplines to software development is productivity. It has been estimated that productivity of producing software has only increased 1 to 2% a year in the last thirty years. Ironically, the computer and its software have contributed significantly to the industry-wide productivity. But, as the old saw goes "The carpenter's house needs repairing and the cobbler's children need shoes," computer professionals have done a poor job of using the computer to do their job.

Engineering disciplines and methodologies are now emerging supported by software tools that address the problems of software development. This paper addresses some of the current software engineering methodologies as a backdrop for the general evaluation of Computer Assisted Software Engineering (CASE) tools from actual installation of and experimentation with some specific tools.

INTRODUCTION

The Engineering Profession is reported to be the second oldest profession of humankind. The first engineering discipline was mechanical engineering to support the building of war machines. Over the years many different engineering disciplines have developed including chemical, electronic, mechanical, civil, etc. to name a few. Common to all engineering disciplines is the use of rigor, modeling, metrics, tools and predefined methodologies. Recently a new engineering discipline has appeared on the scene, called **Software Engineering.** For over thirty years computer software has been being developed and the track record has not been good. Software development projects often miss schedules, are over budget, do not give the user what they want and have defects. One estimate is there are 1 to 3 defects per 1000 lines of deployed code.[1] More and more systems are requiring larger and more complex software for support. As this requirement grows the software development problems grow exponentially. It is believed that software quality can be improved by applying engineering principles. Another compelling reason to bring the engineering discipline to software development is productivity. It is estimated that productivity of producing software has only increased 1 to 2% a year. The average software developer produces 10 to 15 statements a day. This is to be contrasted with the 1960's estimate of 6 to 10 statements a day. It is interesting to note that in the survey work for these statistics, the programming language being used didn't seem to be a factor. This observation has been used by pundits to advocate using the highest level programming language possible.

Software engineering has now emerged as a discipline in its own right. Although its principal base is in computer science, it also makes use of mathematics, psychology, ergonomics, and management science. The software engineer must be able to assess and apply existing computing techniques in a cost-effective and stable way. She or he applies existing knowledge, derived from more fundamental subjects, in the same way as the electrical or mechanical engineer applies physics and mathematics.

This paper addresses some of the current software engineering methodologies as a backdrop to set the stage for the general evaluation of Computer-Aided Software Engineering (CASE) tools and report on the actual installation and experimentation with some specific software engineering tools. For the purposes of this paper, Software Engineering is defined as a strategy for productively producing quality software.

Software Quality

What is quality software? Characteristics of software quality depends on who is viewing the software. Users judge software to be of high quality if it does what they want in a way that is easy to learn and easy to use. Those who design, write and maintain the software also judge the software quality. Software quality characteristics to these include the correct and efficient use of computer resources, and the relative ease with which the system can be designed, coded, tested and maintained. A third view of software quality is from those who must pay for the software systems, i.e. the cost of producing it.

Software Development Processes

Over the years those who produce software systems have developed techniques that follow the engineering approach of decomposing large complex tasks into smaller less complex tasks, often referred to as the Software Development Process or System Development Life Cycle (SDLC), or Software Development Methodology. These methodologies can be represented by several models.

One of the most popular is the Waterfall model. Although there are no agreed to standards for this model, they all contain at least the following activities:

1. Requirements analysis
2. System Design
3. Program Design
4. Program implementation
5. Unit testing
6. Integration testing
7. System testing
8. System delivery
9. Maintenance

Different software development organizations have adopted various combinations of the above processes. Small organizations tend to be relatively informal and the project proceeds from requirements analysis through design and implementation without much fuss. In the larger organizations, however, things are done on a more formal basis. The above activities are defined as project phases with activities and deliverables defined for each phase. Each project must formally follow a rigid set of procedures to proceed from one phase to another. At each phase exit "management" makes one of three decisions.

1. The project is allowed to proceed to the next phase.
2. The project is killed.
3. The project team is sent back to the drawing board (probably the worst case from a project management point of view).

It is within these Software Development Processes that certain methodologies have evolved over time. This evolution and projected future methodologies is illustrated in the following table:

Table 1. SOFTWARE ENGINEERING METHODOLOGIES

MAJOR TASKS	1950 - 1960's	1970 - 1980's	1980 - 1990's	1990 - 2000's
Analysis	None	Victorian novel*	Structured	Object Oriented Analysis (OOA)
Design	None	Structured	Structured	Object Oriented Design (OOD)
Programming	None	Structured	Structured/ Object Oriented Programming (OOP)	OOP/CASE Generated

* Victorian novel means an extremely large narration of the requirements in English in which you had to read the entire document to get an understanding of the analyst's understanding of the user's requirements. Like a Victorian novel, if you didn't read the last page, you had no idea how the story ended.

In structured oriented analysis and design methodologies, primary emphasis is placed on specifying and decomposing system functionality. The object oriented analysis and design approach focuses first on identifying objects from the application domain, then fitting procedures around them. Just 5 to 10% of commercial application developers use structured analysis and design methodologies, according to CASE expert Yourdon. Most all use structured programming. A few leading edge organizations are

moving to object oriented analysis, design, and programming. Indeed, many CASE vendors are joining the object oriented hype.

As the graphical modeling techniques of structured analysis began to spread through systems development organizations in the early 1980's, software engineers began to realize that there was a major problem. The artwork required to create data flow diagrams, entity-relationship diagrams, structure charts, state-transition diagrams, and other graphic models was overwhelming. In addition to the work required to create and maintain the diagrams, classical structured analysis requires a great deal of work to verify the diagrams to ensure that they are complete and consistent. This verification had to be done manually and because it was labor intensive and boring, it tended to be error prone. Consequently, many of the specification errors that should have been found were not.

Many of these problems can be solved with proper automated support; this was well-known when structured analysis was first introduced, but the cost of automation was far higher than most organizations could afford. However, the development of powerful graphics workstations in the mid-1980's led to a whole new industry known as CASE (Computer-Aided Software Engineering).

It is estimated that about 25 to 30% of software engineers have access to personal CASE tools. Many designers, especially those working in real-time and embedded systems, remain skeptical that CASE tools will help them develop better systems. This skepticism can be traced to some of the first CASE tools of 20 years ago. They failed to live up to promises of easing maintenance and ensuring error-free code, making potential purchasers wary. However, it is clearly the way of the future and we can expect that all professional Software Engineers as well as those paying for the software will insist on such tools as time goes on.

The two primary objectives of CASE tools are:

1. Higher Quality Systems
    a. Analysis - Requirements must be:
        1) Correct
        2) Consistent
        3) Complete
        4) Realistic
        5) Needed
        6) Verifiable
        7) Traceable

    b. Design - The Design must be:
        1) A solution to the problem
        2) Modular and well-structured
        3) Portable
        4) Easy for implementers to understand
        5) Well-documented
        6) Cross-referenced with the requirements

    c. Programming - the Programs must be:
        1) Structured
        2) Efficient
        3) Simple
        4) Well-documented
        5) Testable
        6) Easily modified

2. Improved Productivity - Use the power of the computer to improve the productivity of the software engineer and therefore lower the cost of software development projects.

With these objectives in mind, the rest of this paper will give an overview and introduction to CASE tools, their characteristics, uses, etc. It is then followed by an evaluation of various specific CASE tools.

CASE TOOL OVERVIEW[2]

CASE tools provide automated support for carrying out various tasks in the Software Development Life Cycle (SDLC). A CASE tool refers to any software tool that provides the software engineer with automated assistance in the creation, maintenance, or project management of software systems.

CASE toolkits combine CASE tools into an integrated set that serves to automate one particular task of the software development process, such as the design phase. Therefore, toolkits are usually referred to by the particular SDLC phase they support.

CASE workbenches or environments combine CASE tools into an integrated set that serves to automate several tasks across the SDLC. The unique feature of a workbench is the fact that the output of one life cycle phase is automatically carried over and continued into the next phase. A complete CASE workbench supports the following tasks of the software development process:

- Graphically represent design specifications
- Track and cross-reference system information
- Report on system information
- Build prototypes and system simulations
- Generate program code
- Generate programming documentation
- Enforce adherence to standards

CASE Methodology Companions are a special set of integrated CASE tools which serves to automate the software development process according to the rules of a specific structured methodology. Methodology companions can be either a toolkit or a workbench. The tools require the software engineer to adhere to a predetermined series of development steps. Functions within the methodology companion will not allow the developer to skip over a necessary step. In addition, the methodology companion checks the output of a process and confirms that it meets the required specifications and standards.

Some CASE toolkits are designed to work under the direction of a specified structured methodology. For example, the Analyst/Designer toolkit from Yourdon, Inc. only supports the Yourdon Structured Design methodology, while most recently developed CASE tools support multiple methodologies such as System Architect from Popkin Software & Systems, Inc. Some of the most popular methodologies are:

* **Yourdon/DeMarco** - Structured methodology

* **Gane/Sarson** - Structured methodology

* **Ward/Mellor** - Real time extensions to the Structured methodology

* **Rumbaugh[3]** - Object oriented methodology using the Object Modeling Technique (OMT)

* **Shlaer/Mellor** - Object oriented methodology

* **Information Engineering**[4] - Methodology introduced by Clive Finklestein and later popularized by James Martin. The methodology addresses the lack of cohesiveness between the different phases of other SDLC's. It addresses strategic planning, as well as analysis and design, and is data driven. This focus allows the developer to concentrate the model building activity on the more stable facet of an organization, its data. The key to the power of Information Engineering is its degree of integration. While many of the techniques used in Information Engineering originated elsewhere, Information Engineering clearly defines how deliverables from one technique relate to deliverables from other techniques within and across development stages.

* **SSADM**[5] - Structured Systems Analysis and Design Methodology is an enhanced rewrite of the Learmonth and Burchett Management Systems System Development Methodology (LBMS-SDM) which is a highly structured, data driven method of systems analysis and design that covers all stages of system development. The input to the methodology is an initial study report and the outputs are stated as program specifications, user procedures, operating instructions and the file design or data base schema. It incorporates highly defined procedures for both logical and physical design, and relies heavily on the data analysis approach. The major techniques used in the methodology are data flow diagrams, logical data structuring technique, third normal form data analysis, a data dictionary, walkthroughs, and reviews. These techniques are based on the work of DeMarco, Yourdon, Gane and Sarson, Martin, Codd, and Bachman.

* **DOD-STD-2167A**[6] - Defense System Software Development, establishes a standard for planning and controlling software development on large software development projects. The standard is primarily used within the United States for engineering-oriented system acquirers (e.g., NASA, Department of Transportation, Department of Defense) and contractors. Some countries outside the US use the standard too, generally for work under contract with a US firm developing a system in that country.

US industry and Government personnel worked for several years to establish a uniform standard for large software development projects. The US Government released an initial version of the standard, DOD-STD-2167, on June 4, 1985 and then on February 29, 1988 a revision, DOD-STD-2167A, was released. The revision resulted in a publication that was shorter and simpler. The earlier 2167 standard presented the classical "waterfall" model. Whereas 2167A is development cycle independent and software method independent. The contractor defines how to organize activities over time, and includes the description in the Software Development Plan. This means that the contractor can select his process model -- waterfall, rapid prototype, multiple build, spiral, or some hybrid of these. 2167A allows the contractor to choose his software method, whether it be object oriented or structured or some in-house version, and it too is documented in the Software Development Plan.

Integrated CASE Environment is a suite of linked tools that addresses different aspects of the development process. Typically, these tools are produced by different vendors. Tool integration is intended to produce complete environments that support the entire software development process. There are four types of integration:

1. Presentation integration --sharing a look and feel from the user's perspective.

2. Data integration --sharing data among tools and managing the relationships among data objects produced by different tools.

3. Control integration --event notification between tools and the ability to activate tools under a program's control.

4. Platform integration --inter-operability via distributed processing which make it possible to use network-based file systems and network communication to convert and transmit files from one execution environment to another.

Software development tools are classified as **"vertical"** - tools that are used for a specific phase of the process- and **"horizontal"** - tools that are used throughout the software development process. In the following figure the various tools are depicted in an overall framework that shows a model of an integrated tool environment:

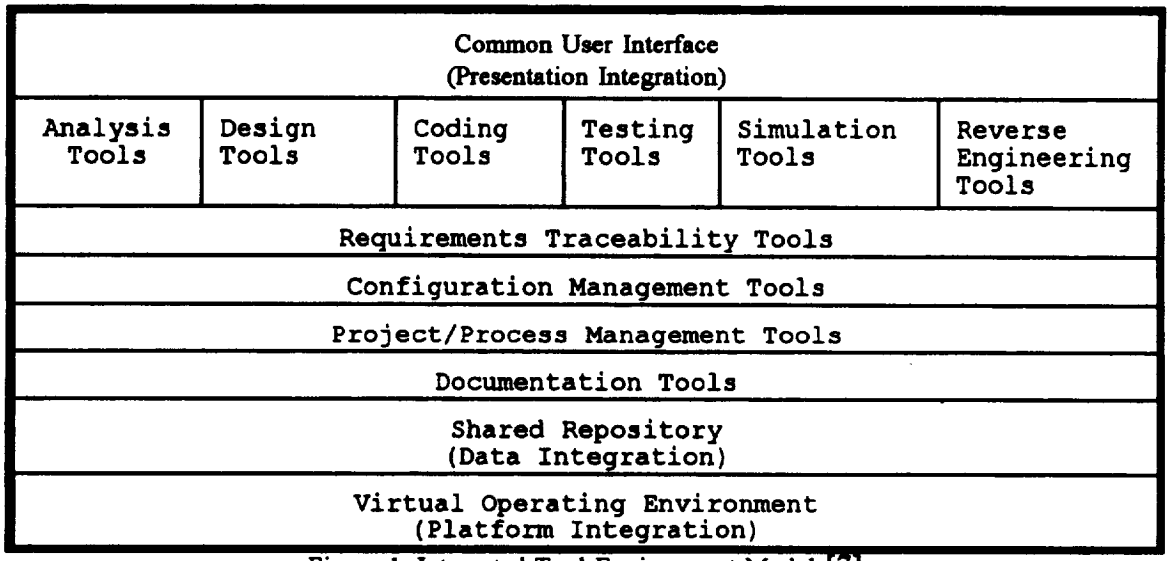


Figure 1. Integrated Tool Environment Model [7]

Vertical CASE tools assist software engineers by automating the tasks specific to a phase of the software development life cycle. Vertical CASE tools support the automatic creation of the following:

* **Data Flow Diagrams**(DFD). Software engineers use data flow diagrams for structured systems analysis and object oriented analysis. Data flow diagrams allow a user to model the system from a "functional" perspective. They do this by first defining the system at the top level to show its overall purpose and how it interacts with external objects. Then, the system is divided into individual components or processes using decomposition techniques. Every process transforms the incoming data, then passes its output data to another process or to a data store. For primitive processes that are not decomposed, the user generates a process specification ("mini-spec") from a user-modifiable template. In integrated workbenches, this information is stored in a central repository and lays the groundwork for design and implementation. The purpose of the Data Flow Diagram is not to specify, but to declare. It declares component processes that make up the whole, and it declares interfaces among the components. Although the DFD is the primary emphasis in structured analysis it has less emphasis in object oriented analysis.

* **Hierarchical Data Structure Diagrams (HDSD)**. Software engineers use hierarchical data structure diagrams during the design phase to decompose data flows and data stores that are defined in data flow diagrams and structure charts.

* **Entity Relationship Diagrams**(ERD). Entity relationship diagrams were added to the structured analysis method by Yourdon to model the system's stored data layout at a high level of abstraction. Yourdon called this modification, "modern" structured analysis. It highlights relationships between data stores on the DFD that would otherwise be seen only in the process specification. ERDs are also used during the design phase to model system data by defining entities and their relationships, particularly when the application is data-intensive.

* **Control Flow Diagrams**. Control flow diagrams are used to model the event-driven aspects of real-time systems. Control flow diagrams show the interaction between discrete-valued control signals and processes. They describe the variety of pathways that the program can take depending on external events.

* **Finite or State Transition Diagrams**. Finite or state transition diagrams are used to model the behavior of event-driven systems. State transition diagrams allow users to view the system from an "event-control" perspective. They show the conditions and actions that cause a system to change from one state to another.

* **Structure Charts**. System designers use structure charts for structured design. Structure charts allow users to define the software architecture, including the interfaces between modules. Modules are defined hierarchically by specifying which modules "call" other modules.

* **Object Diagrams**. The Object Diagram is typically the first diagram built using the Object Modeling Technique of the Rumbaugh Object Oriented Analysis methodology. This diagram contains analysis and design information about how objects in the system are related to one another. It is a graph whose nodes are object classes and whose arcs are relationships among classes.

Horizontal CASE tools used throughout the software development life cycle are described as follows:

* **Shared Repository or Data Dictionary.** The data dictionary (also called Encyclopedia by KnowledgeWare's Application Development Workbench) is a repository that contains entries for all data elements, objects, etc. that are pertinent to the system, with precise, rigorous definitions so both the user and analyst will have a common understanding of all inputs, outputs, components of stores, and intermediate calculations. The Data Dictionary is used by CASE tools for centralized storage of intelligent information about the system. Information about the system is collected in this repository throughout the development life cycle allowing ease of transition among the planning, analysis, design and construction phases of the life cycle. This repository is also constructed with reverse engineering tools allowing the automatic creation of tools to assist with system understanding and forward engineering.

* **Requirements Traceability.** Requirements Traceability tools produce reports to enable software developers to track original customer requirements throughout the early phases of the software development process. Tracing customer requirements enhances management control and improves communication between software developers and their customers.

Examples of Requirements Traceability reports include:

- ☐ objects satisfying a particular requirement
- ☐ structure chart modules traced back to data flow processes
- ☐ data flow processes linked to structure chart modules

* **Reverse Engineering.** Although shown in the above chart as a vertical tool because it is used after code is produced, reverse engineering tools can be used to support the entire life cycle. It is this view that the Cadre tool Ensemble has. Perhaps the title of Software Rejuvenation would be more fitting. Software Rejuvenation addresses the challenge of software system maintenance by trying to increase the overall quality of an existing system. It looks back at the workproducts of a system to try to derive additional information or to reformat them in a more understandable way. There are several aspects of software rejuvenation to consider, including:

- Redocumentation which is the static analysis of source code whose product is additional information that assists maintainers in understanding and referencing the code. The analysis does nothing to transform the actual code; it merely derives information.

- Restructuring which actually changes the code by transforming ill-structured code into well-structured code.

- Reverse engineering which looks back from the source code to the products that preceded it, re-creating design and specification information from the code.

- Re-engineering which is broader still, is where the existing system is reverse engineered and then "forward engineered" to make changes to the specification and design to complete the logical model; a revised system is then generated from the new specification and design.

Interestingly, NASA is very interested in this area of CASE tools not necessarily for the maintenance of systems but for the evaluation of systems produced by contractors. Areas where these tools could be helpful to NASA would be:

- Requirements traceability. The tools could be used to reverse engineer the delivered code back to the design specifications. The tool produced specifications could be compared to the original design specifications and differences explained, justified, or corrected. Using special requirements traceability tools which span the entire life cycle, design specifications can be traced back to the requirements, given that the CASE tool was used to produce the original requirements.

- Quality Assurance. The tools could be used to analyze the delivered code, generate comprehensive test cases, validate that all important paths were tested, generate system metrics, etc.

Also, interestingly, these tools can be used by the contractors to generate the required documentation. It is possible to reverse engineer a system of source code back to design specifications to include items such as structure charts, control flow diagrams, cross-references, library calls, etc. Then using special documentation tools with DOD-STD-2167A templates generate the required design documentation. If the CASE tool was used throughout the project's life cycle, requirements through code generation, CASE documentation tools are able to extract, assemble and format total system documentation . Some CASE vendors see this as a software engineer productivity tool. "No more having to do that awful program documentation." Let the tool do it.

CASE TOOL SELECTION CRITERIA

Over the past decade a broad array of software tools has been introduced to the software engineering marketplace. Software tools that support analysis, design, testing, maintenance activities, configuration management, and even quality assurance combine to create a CASE environment. Unfortunately,

vendor claims for specific tools and the performance of these tools in actual practice are often dramatically different. For this reason, it is very important to assess tools carefully using the criteria in [8] as a guide.

A Case Tool Checklist

The above referenced generic tools checklist can be used effectively when CASE is considered. However, CASE systems have a set of specific attributes that should be considered as well. A specialized checklist for CASE selection can be found in [9]. Probably no existing CASE system will exhibit all of the attributes listed in the above reference. The best systems offer a reasonable subset. Use the attributes checklist as a guide in evaluating vendor offerings.

A COMPARATIVE ANALYSIS AND EXPERIENCES WITH SOME SPECIFIC CASE TOOLS

As a Summer Faculty Fellow working for NASA, the author had the opportunity to do research in the CASE tools area. The research involved reading the many texts that there never seem to be enough time for back at the university, discussing current CASE issues with fellow Faculty Fellows, attending CASE product seminars and demonstrations and getting some actual hands-on experimentation with real CASE tools. The author is grateful to the companies that loaned evaluation copies for this research. The following is a report on the results of experimenting with each tool.

Software Through Pictures (StP), C Development Environment from Interactive Development Environment is written for the UNIX environment and was installed on a Sun Sparc 2 running in a Sun OS 4.1.2 environment.

StP is an integrated, open environment, built upon a multi-user, object-based repository that facilitates the integration of StP tools with other development tools in a consistent, seamless manner. StP provides open and published interfaces including diagram file formats and repository schema, as well as user-modifiable configuration and message files. The StP product includes a family of graphical editors that implement several commonly used analysis and design methods, as well as code generators. For example, there are analysis tools that support the structured analysis using either the notation developed by Yourdon/DeMarco and the notation developed by Gane/Sarson. It also supports modeling of real-time systems. StP products provide user-modifiable templates for object annotation, documentation preparation, code generation, SQL schema generation, and program design language.

Paradigm Plus, from Protosoft Inc. , is a configurable CASE tool that uses an object-oriented information model to provide support to a wide range of software engineering activities practiced throughout the application development life cycle.[10] Paradigm Plus comes with a set of application development rules and procedures including:

- Rumbaugh OMT - Object Modeling Technique
- Rumbaugh 93 - latest extensions to OMT
- Booch OOD - Object Oriented Design
- HP Fusion - Object Oriented Analysis and Design Method
- ProtoSoft OOAD - Object Oriented Analysis/Design
- EVB - Ada Object Oriented Design
- others

The set of application development rules and procedures are implemented by using the Paradigm Plus

CASE Development Kit (CDK) technology. The CDK *Paradigm Definition Language* (PDL) provides the ability to specify a particular set of rules and procedures to include:

- Meta Model
- Rules and Semantics
- Diagram Notations
- Menus
- Toolboxes
- Icons
- Dialog Boxes
- Hypertext Help
- Standard Reports
- Code Generation

Paradigm Plus incorporates many features that set it apart from conventional CASE tools. It is Method Independent, allowing a software engineer to choose the method that is most suited to a given project. It is language independent, allowing the software engineer to specify a neutral definition of the software design, and have it automatically mapped to language specific constructs at code generation time. Paradigm Plus is platform independent, being available on Windows, Solaris 1.x and 2.x, HP PA-RISC and IBM RS/6000. Paradigm Plus stores all information about a project in an integrated, active Object Repository.

Teamwork, Ensemble [11], from Cadre Technologies, Inc. is written for both the OS/2 and the UNIX environment. The Teamwork product set is comprised of a number of integrated development tools that support the full software development life cycle. All tools support multi-user, multi-windowed, multi-tasking environments. All Teamwork tools share a common repository for each project, known as the Teamwork Project Database, which serves as a repository for data dictionaries, diagrams, process specifications, and project management data. The Teamwork series includes Teamwork/IM, Teamwork/SA, and Teamwork/RT, systems analysis tools which support information modeling, systems analysis, and real-time analysis, respectively. Other Teamwork analysis/design tools include Teamwork/SD, which supports structured design, Teamwork/Ada, a design, navigation and documentation tool for Ada projects, and Teamwork/OOD, which provides direct support for the object-oriented design (Schlaer/Mellor).

Teamwork supports code generation via their Ada and C Source Code builders. Tool sharing and portability is provided by Teamwork/IPSE which enables all Teamwork products to integrate with other software applications. A new member of the Teamwork family is Teamwork/SIM, which integrates dynamic modeling and real-time simulation into the Teamwork environment, providing verification of real-time system behavior.

Ensemble is a modular tool suite that automates software development and maintenance for C professionals. Integrating a wide range of functions required for understanding, constructing, testing, and documenting code under a graphical user interface and shared database, Ensemble complements existing tools by fitting easily into current development environments and enabling a phased adoption of its modules. For a detailed product description, see CADRE's teamwork Product Overview[12].

EVALUATIONS/RECOMMENDATIONS

1. Product Comparisons and Comments. Because of limited time and hardware availability, the author was not able to explore in-depth the Software through Pictures (StP) and the Paradigm Plus products. With that in mind, the following are some general comments about the products.

The **StP** product was an Upper CASE tool. I experimented with the graphical editors normally used in the Requirements phase to draw Data Flow Diagrams, Entity Relationship Diagrams, etc. I found them

a bit cumbersome to use. You definitely need a strong Unix System Administrator to install the package. IDE has recently announced a Reverse Engineering tool and wanted to make it available for our evaluation. However, time and IDE policy prohibited this. IDE product documentation was excellent. The marketing literature is also excellent. IDE is selling I-CASE, but we only know of an Upper CASE tool and the recently announced Reverse Engineering tool.

The **Paradigm Plus** product supports OOA (Object Oriented Analysis) using the OMT (Rumbaugh) methodology, both of which are new and seem very complex to the author. The product installed without any problems. The version we tried runs in a PCDOS/Windows environment. We had draft documentation for a new up coming release. We found the product easy to use with the only negative aspect being that charts printed on the printer were low quality. The lines and text were too thick. We were unable to tell if this could be controlled by some parameters. Although the sales literature lead you to believe that Paradigm Plus will generate code, no indication of this function could be found in our evaluation copy literature.

CADRE's TEAMWORK/ENSEMBLE was the product that most time was spent examining. Most of the time was spent with the reverse engineering tool ENSEMBLE. We found that the complete package was indeed an I-CASE tool. A CADRE marketing representative installed and demonstrated the product. The only trouble we had was installing a new key file. This is a file used by the system to control the licensing of the products. Checksums are generated from the information in the key file and compared by predefined algorithm. Therefore, the data in the keyfile has to be "exactly" correct. The first time, after many hours with CADRE's help-desk (hot-line), they had to email us a file. We were more successful the second time when CADRE gave us an extension on our evaluation time.

We found their people on the hot-line to be outstandingly helpful. They were courteous, prompt, and guided us through several problems. They also followed up to ensure we had indeed solved our problem.

We reversed engineered two programs with very satisfactory results. One was the huge C program, Dynamic Environment Communications Analysis Testbed (DECAT), the other a modest C program written by another faculty fellow. We encountered no problems and produced some useful documentation from the reverse engineered repository.

Another faculty fellow experimented with the Upper CASE tool to design and generate an Ada program. This effort is still underway and will be completed after the author leaves.

2. Recommendations

My recommendations for the Systems Design Section of the System Engineering Branch of the Tracking and Communication Division at NASA JSC regarding CASE tools fall into two categories, internally developed and maintained systems and externally developed and maintained systems.

a. Internally developed and maintained systems. First of all I recommend you appoint someone to be your software engineering expert. Then, you must adopt a methodology. The gurus agree unanimously that a methodology is an absolute requirement for software success. Ed Yourdon says "It is better to pick any methodology than it is to pick mine."[13] Gerald Weinberg states "Bringing in a CASE tool where there is no methodology, where anybody can build systems any way they want, is guaranteed to be a waste of time." He goes on to say "A methodology is a way of thinking, not a substitute for it." The learning curve for methodologies is typically from 13 to 24 months.[14] Because you will need

some experience and knowledge of methodologies and CASE tools to manage the externally developed and maintained systems, this will prepare you for managing the external projects better. Although the author is not totally up to speed on OOA and OOD, from the research done so far indications are that the organization should be adopting one of the Object Oriented Methodologies. As Object Oriented Programming brings promises of improved programming concepts such as code reusability, OOA promises similar benefits that are described in the following reference.[15]

After choosing a particular methodology, then a CASE tool should be purchased. The tool should be an I-CASE tool that includes reverse engineering. I recommend that CADRE's Teamwork/Ensemble package be seriously considered. An approximate license cost is $70,000 for a single user plus an annual maintenance. Official price quotes should be obtained directly from CADRE. CASE tools, especially I-CASE are large complex systems. They are not intuitive for a simple reason. They automate complicated design techniques. (The most advanced word processing software will not make you a good author.) Therefore, I also recommend that your software engineering expert obtain detail knowledgable on CASE tools by attended formal training for both the methodology and the CASE tool. Using the tool/methodology will prepare your organization for dealing with the following:

b. Externally developed and maintained systems. The real issue with recommendations in this area is the DOD-STD 2167A, Defense System Software Development standard currently used by NASA. According to Coad/Yourdon[16] 2167A is model and method free. The contractor is allowed to select his process model - waterfall, rapid prototype, etc, and also select his method - structured, object, etc. However, this selection must be agreed upon by the Government. Therefore, it would seem that NASA could specify the model/method and CASE tools to be used by the contractor. The author recommends that NASA use the waterfall model and object oriented methodology. For really new technologies a variation of the waterfall model could be used where prototyping was used to gather requirements. For project management reasons, the author is opposed to the pure prototyping models. For a detailed mapping of OOA to DOD-STD-2167A, see the above reference.

Reverse engineering tools seem to be model/method independent, and only source program language independent. Since most programs written for NASA is written in either Ada or C, a reverse engineering tool could be used by NASA to gather metrics, verify "hot" spots for testing and even generate acceptance test cases. Again, CADRE's Ensemble tool would be very useful for these purposes.

If NASA could specify to the contractor the model/method/I-CASE then additional benefits could be gained such as the presumed increased contractor software engineering productivity, requirements traceability, software and analysis reuse, increased system reliability, etc. It is strongly recommended that NASA set this specification as a near term goal.

REFERENCES

[1] Van Tyle, Sherrie. Electronic Design, May 13, 1993.

[2] SRA: Using CASE Methodologies. Science Research Associates, 1991.

[3] Rumbaugh, James, Object-Oriented Modeling and Design, Prentice-Hall, 1991.

[4] Madison, R. N., Information System Methodologies, Wiley Hayden, 1983.

[5] ibid.

[6] Coad/Yourdon, Object-Oriented Analysis, second edition, 1991

[7] IDE Product Catalog 90-00713-492-30M, Interactive Development Environments. IDE, 1992.

[8] Pressman, R. S., Making Software Engineering Happen. Prentice-Hall, 1988.

[9] ibid.

[10] Paradigm Plus User's Guide, Revision 1.1, June 1993, first edition.

[11] CASE Trends, Volume 5, Number 4, Summer 1993.

[12] CADRE's teamwork Environment Reference Manual, Release 4.0, 1990.

[13] CASE Trends, Volume 5, Number 5, July 1993, pp56.

[14] ibid

[15] Coad/Yourdon, Object-Oriented Analysis, second edition, 1991

[16] ibid, Appendix B

N94- 25383

SPACE STATION FLEXIBLE DYNAMICS UNDER PLUME IMPINGEMENT

Final Report

NASA/ASEE Summer Faculty Fellowship Program - 1993

Johnson Space Center

Prepared By:	Trevor Williams, Ph.D.
Academic Rank:	Associate Professor
University & Department:	University of Cincinnati Department of Aerospace Engineering Cincinnati, Ohio 45221

NASA/JSC	
Directorate:	Engineering
Division:	Guidance, Navigation & Aeronautics
Branch:	Control & Guidance Systems
JSC Colleague:	John Sunkel, Ph.D.
Date Submitted:	September 14, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

Assembly of the Space Station requires numerous construction flights by the Space Shuttle. A particularly challenging problem is that of control of each intermediate station configuration when the Shuttle Orbiter is approaching it to deliver the next component. The necessary braking maneuvers cause Orbiter thruster plumes to impinge on the station, especially its solar arrays. This in turn causes both overall attitude errors and excitation of flexible-body vibration modes. These plume loads are predicted to lead to CMG saturation during the approach of the Orbiter to the SC-5 station configuration, necessitating the use of the station RCS jets for desaturation. They are also expected to lead to significant excitation of solar array vibrations.

It is therefore of great practical importance to investigate the effects of plume loads on the flexible dynamics of station configuration SC-5 as accurately as possible. However, this system possesses a great many flexible modes (89 below 5 rad/s), making analysis time-consuming and complicated. Model reduction techniques can be used to overcome this problem, reducing the system model to one which retains only the significant dynamics, i.e. those which are strongly excited by the control inputs or plume disturbance forces and which strongly couple with the measured outputs. The particular technique to be used in this study is the *subsystem balancing* approach which was previously developed by the present investigator. This method is very efficient computationally. Furthermore, it gives accurate results even for the difficult case where the structure has many closed-spaced natural frequencies, when standard modal truncation can give misleading results. Station configuration SC-5 is a good example of such a structure.

INTRODUCTION

Assembly of the Space Station requires numerous construction flights by the Space Shuttle, regardless of the details of the station configuration that is finally selected. A particularly challenging problem is that of control of each intermediate station configuration when the Shuttle Orbiter is approaching it to deliver the next component. The necessary braking maneuvers cause Orbiter thruster plumes to impinge on the station, especially its solar arrays. This in turn causes both overall attitude errors and excitation of flexible-body vibration modes. These plume loads are predicted to lead to CMG saturation during the approach of the Orbiter to the SC-5 Space Station Freedom (SSF) configuration, necessitating the use of the station RCS jets for desaturation. They are also expected to lead to significant excitation of solar array vibrations.

It is therefore of great practical importance to investigate the effects of plume loads on the flexible dynamics of station configuration SC-5 as accurately as possible. However, this system possesses a great many flexible modes (89 below 5 rad/s), making analysis time-consuming and complicated. Model reduction techniques can be used to overcome this problem, reducing the system model to one which retains only the significant dynamics, i.e. those which are strongly excited by the control inputs or plume disturbance forces and which strongly couple with the measured outputs. A considerable amount of work has been carried out on the topic of model reduction in the past. Well-known methods include *modal truncation* [1], based either on the natural frequencies of the structure or its modal costs, and *balancing* [2] of the entire structure and then truncation to retain a dominant model for it. An advantage of the balancing approach is that it typically yields a more accurate reduced-order model than does simple modal truncation. This is particularly true when the structure possesses clustered natural frequencies, as is often the case for realistic flexible space structures. However, the disadvantages of balancing are its high computational cost, possible numerical sensitivity problems resulting from the large matrices being operated on, and the difficulty involved in providing a physical interpretation for the resulting balanced "modes".

The purpose of the work reported here is to investigate the performance of the alternative *subsystem balancing* technique when used to study the plume impingement problem on SC-5. This method, introduced in [3][4], further developed in [5], and then applied to a simplified SSF model in [12], retains the desirable properties of standard balancing while overcoming the three difficulties listed above. This is achieved by first decomposing the structural model into subsystems of highly correlated modes, based on the *modal correlation coefficients* derived in [4] from the controllability and observability Grammian matrices [6] of the structure. Each subsystem is approximately independent of all others, so balancing each separately and concatenating the dominant reduced-order models obtained yields roughly the same result as balancing the entire structure directly. The computational cost reduction produced by this subsystem technique is considerable: an operation count reduction by a factor of roughly r^2 if the system decomposes into r equal subsystems. The numerical accuracy of the resulting reduced-order model is also improved considerably, as the matrices being operated on are of reduced dimension; this avoids the numerical conditioning problems noted in [8][9] for standard balancing. Furthermore, the modes of the reduced model do now permit a clear physical interpretation. This is a consequence of the fact that each correlated subsystem must necessarily only include modes with close natural frequencies. The balanced modes of each subsystem are therefore, to first order, linear combinations of repeated-frequency

modes, and so can themselves be taken as an equally valid set of physical modes. Balancing the entire structure, on the other hand, combines modes of widely differing frequencies, making interpretation difficult.

The graphs and tabulated results given in this report demonstrate the improvements achievable by using the subsystem balancing technique for model reduction. In fact, the original 94-mode model for the SC-5 SSF configuration was reducible to a 10-mode model at the cost of a modeling error of less than 2%. This then allows analysis and simulation of the effects of plume loading to be carried out much more efficiently than if the entire model were used.

THEORETICAL BACKGROUND

Consider an n-mode model for the structural dynamics of a modally damped, non-gyroscopic, non-circulatory FSS with m actuators and p sensors, not necessarily collocated. This model can be written in modal form [1] as

$$\begin{aligned} &\ddot{\boldsymbol{\eta}} + diag(2\zeta_i\omega_i)\dot{\boldsymbol{\eta}} + diag(\omega_i^2)\boldsymbol{\eta} = \hat{B}\mathbf{u}, \\ &\mathbf{y} = \hat{C}_r\dot{\boldsymbol{\eta}} + \hat{C}_d\boldsymbol{\eta}, \end{aligned} \tag{1}$$

where $\boldsymbol{\eta}$ is the vector of modal coordinates, $\mathbf{u}$ that of applied actuator inputs and $\mathbf{y}$ that of sensor outputs, and ω_i and ζ_i are the natural frequency and damping ratio of the i^{th} mode, respectively. For the typical FSS [7], the $\{\zeta_i\}$ are quite low (e.g. 0.5 %), and the $\{\omega_i\}$ occur in clusters of repeated, or nearly repeated, frequencies as a result of structural symmetry.

Defining the state vector $\mathbf{x} = (\dot{\eta}_1, \omega_1\eta_1, \cdots, \dot{\eta}_n, \omega_n\eta_n)^T$ for this structure yields the state space representation $\dot{\mathbf{x}} = A\mathbf{x} + B\mathbf{u}$, $\mathbf{y} = C\mathbf{x}$, where $A = blkdiag(A)_i$, $B = (B_1^T, \cdots, B_n^T)^T$ and $C = (C_1, \cdots, C_n)$, with

$$A_i = \begin{pmatrix} -2\zeta_i\omega_i & -\omega_i \\ \omega_i & 0 \end{pmatrix}, B_i = \begin{pmatrix} \mathbf{b}_i \\ 0 \end{pmatrix} \text{and } C_i = \begin{pmatrix} \mathbf{c}_{ri} & \mathbf{c}_{di} / \omega_i \end{pmatrix}; \tag{2}$$

$\mathbf{b}_i$ is the i^{th} row of $\hat{B}$, and $\mathbf{c}_{ri}$ and $\mathbf{c}_{di}$ are the i^{th} columns of $\hat{C}_r$ and $\hat{C}_d$, respectively.

The problem studied here is that of obtaining a reduced-order model

$$\begin{aligned} &\dot{\mathbf{x}}_r = A_r\mathbf{x}_r + B_r\mathbf{u}, \\ &\mathbf{y}_r = C_r\mathbf{x}_r \end{aligned} \tag{3}$$

for this structure for which the normalized output error

$$\delta^2 = \int \|\mathbf{y}(t) - \mathbf{y}_r(t)\|_2^2 dt \Big/ \int \|\mathbf{y}(t)\|_2^2 dt \tag{4}$$

is acceptably small. Of course, the size of δ will depend on the order, n_r, chosen for the reduced model. A good model reduction procedure should ideally provide information allowing an intelligent choice for n_r to be made so as to achieve a specified δ value.

Two techniques for model reduction that have been extensively studied are those of modal truncation and internal balancing. The new method implemented in this report, subsystem balancing, can be regarded as a generalization of the two established techniques; it includes both as special cases, but permits a wide range of other solutions also. It is this additional freedom that allows subsystem balancing to produce superior results to those obtainable by means of the traditional methods.

Model reduction by subsystem balancing proceeds by first dividing the given structure into subsystems of highly correlated modes. The dimensions of these subsystems are determined by the choice of a *correlation coefficient threshold* value by the analyst. This quantity is used as follows. Two modes which have nearly equal natural frequencies and non-orthogonal mode shapes will be approximately in resonance, and so can strongly excite each other. An input which excites the first mode will thus typically also indirectly excite the second. It is therefore important to treat the two modes as a single unit, or subsystem, when considering whether or not to retain them in a reduced-order model. Such a pair of close modes can be shown to have a high correlation coefficient; the basis of subsystem balancing is therefore to define subsystems made up of all modes which have correlation coefficients greater than the defined threshold value. It can be seen that selecting a threshold of 0 implies that all modes are deemed to be correlated, i.e. there will be a single "subsystem" which is in fact the entire system. This is equivalent to the case of standard balancing. Conversely, choosing a threshold of 1 leads to all modes being deemed to be uncorrelated, and so dealt with in isolation: this amounts to the modal truncation technique.

Each of the subsystems obtained for the chosen threshold value is then balanced independently, and a reduced-order model for it generated by deleting all balanced states corresponding to *Hankel singular values* [2] below some specified threshold. Note that the singular value weighting described in [10] could be applied, if desired, without changing the argument in any way. Similarly, frequency weighting of the Hankel singular values can easily be incorporated to deal with input signals which have a known frequency spectrum. This is actually done in the present application, where the inputs are steps (representing thruster firings) rather than the impulses classically considered in model reduction problems. The resulting reduced-order subsystem models so obtained are then combined to yield a dominant, approximately balanced, reduced-order model for the full system.

USER INTERFACE

This section describes the user interface to the model reduction package which was developed as part of this contract. This software consists of a library of Matlab m-functions, with mrmain3 calling all the other functions internally. The package is installed on the Sun SparcStation 2 *carmel* in the Interactive Analysis Laboratory in Building 16, and has also been provided in a Macintosh version to personnel at JSC and Dynacs Engineering Co. The information that follows details the user interface for mrmain3. This function, together with the purpose-built second level functions that it calls, have

extensive in-line documentation, facilitating future use and/or modification. Listings of all these functions are provided in the report [12].

Input arguments

om: The natural frequencies (rad/s) of the structure, input as either a row or column vector. Any rigid-body modes must precede the flexible modes and be represented by hard zero frequencies.

phia: The influence matrix, in mass-normalized coordinates, corresponding to the specified actuator locations. If the structure has n modes and m actuators, *phia* will be an (n x m) matrix.

phis: Similar to *phia*, but for sensor stations or positions of outputs of interest (e.g. solar array tips).

Prompts to user

The following prompts are generated by mrmain3 when running. They allow each run to be customized as desired by the user.

Output the time taken for each step?: The time required for each matrix decomposition, etc., is output to the screen if requested. This allows the progress of the model reduction procedure to be monitored, as well as giving an indication of which steps are the most computationally intensive.

Vectorize? (Faster, but requires more storage): In Matlab, for loops are typically an order of magnitude slower to execute than the equivalent "vectorized" operation. For instance, s=0; for i=1:n, s=s+x(i); end; runs considerably slower than does s=x*ones(n,1). If vectorization is requested, computation of the system Grammian matrices and correlation coefficients is put into the form of vector-matrix operations rather than loops; this is indeed considerably faster, but requires some additional temporary storage arrays.

Structural damping ratio, % (default is 0.5%): The specified damping ratio is applied uniformly to all flexible modes of the full structural model.

Print frequencies in Hz?: The mean frequency of each subsystem can be output in either rad/s or Hz, as desired.

Desired controllability threshold?: This threshold value is used to determine which modes are correlated in a controllability sense. The system is then broken down into disjoint sets of modes (subsystems), where modes with a controllability correlation coefficient greater than the specified threshold are deemed to be correlated. Taking a threshold value of 0 implies that all modes are considered correlated, i.e. the method reduces to standard balancing. Conversely, a threshold value of 1 implies that no modes are taken together: this is modal truncation. Intermediate values allow the dimensions of the resulting subsystems to be specified to a large extent; reducing the threshold reduces the number of subsystems, so increasing their dimension. The program allows the user to try various different values for this threshold until he is satisfied with the resulting set of

subsystems. At this point, a negative value is input to cause the program to proceed to the next step.

Desired overall threshold?: This threshold is used in a similar fashion to the controllability threshold, but both controllability and observability are now taken into account. This yields the final subsystem distribution output by the program (in *modemap*) and used to obtain the reduced-order model.

Compare step responses?: If requested, the step responses of the full and reduced-order models are computed, plotted, and the relative differences (i.e. reduced-order model error) output for each input-output channel.

Compare frequency responses?: If requested, the Bode amplitude plots of the full and reduced-order models are similarly computed and plotted, and the relative differences (i.e. reduced-order model error) output for each input-output channel.

Desired truncation measure?: Two types of measure can be used to define the number of modes retained in the reduced-order model. If a positive integer is input, this is taken to be precisely the desired reduced model order. On the other hand, if a real number in the interval [0, 1) is input, this is taken to be the desired relative error in the reduced-order model step response, and the model order required to achieve this is computed. (Note that this second option is only an approximation, and should be treated as such.) The program loops over this step, allowing the user to try a sequence of various different reduced model orders to determine which one is most satisfactory for his purposes. This is quite a rapid operation when using mrmain3, as nearly all computations need only be done once. Essentially the only step that must be repeated when trying a new model order is that of calculating the step and/or Bode plots of the reduced-order model.

Output arguments

am, bm, cm: The reduced-order state-space model matrices $\{A_r, B_r, C_r\}$ obtained.

modemap: This matrix specifies which physical modes are grouped into which subsystems in the decomposition based on overall correlation coefficients. In particular, the i^{th} column of *modemap* lists the modes making up the i^{th} of these subsystems.

DISCUSSION OF RESULTS

Results will now be provided which illustrate the behavior of the subsystem balancing technique when applied to a structural model [11] of the SC-5 configuration of Space Station Freedom. This structure possesses light damping (estimated to be 0.5% of critical), and a large number of closely-spaced vibration modes (94 flexible modes in the model considered, of which 89 are below 5 rad/s). The model has 12 inputs: 6 Reaction Control System (RCS) thrusters, and 6 disturbance inputs which represent the lumped forces and torques about all three axes that result from Shuttle Orbiter plume impingement during approach. Note that the Control Moment Gyroscopes (CMGs) that would also be used on SC-5 do not need to be considered as inputs here: this is because CMGs incorporate torque shaping techniques so as to not significantly excite flexible vibration modes. The measured outputs are the 3 angular rates sensed by the rate gyros

on the station. (The movements at other positions of interest, for instance the solar array tips, could also be considered if desired; the method remains exactly the same.)

The three numerical parameters which the user must select when running mrmain3 are: the controllability coefficient threshold value; the overall (controllability and observability combined) coefficient threshold; and the order of the final reduced-order model. It should be noted that, as previously pointed out, the program allows the user to try a sequence of values for each of these quantities until satisfactory results are obtained. The effects of different choices for these three parameters will now be examined.

Figure 1 illustrates the role of the controllability threshold coefficient ρ_{ct} in determining the subsystem decomposition. The two solid graphs give the maximum and minimum subsystem dimensions that result for values of ρ_{ct} ranging between 0 (standard balancing) and 1 (modal truncation). The dashed graph then shows the number of subsystems obtained for each threshold value. It can be seen that the system decomposition does indeed change, as expected, from that of standard balancing (the only subsystem is the entire model, with 94 modes) to that of modal truncation (there are 94 single-mode subsystems) as the threshold increases from 0 to 1. It can be noted that the evolution of subsystem dimensions is somewhat discontinuous: for instance, large changes occur for thresholds between 0.05 and 0.15, whereas there are hardly any differences between 0.55 and 0.70. A consequence of this is that it is not always possible to find a threshold value which will yield a particular maximum subsystem order. However, it *is* always possible to obtain a good working value which gives a totally acceptable subsystem partition. From practical experience, it has been observed that a good choice of threshold value is generally one which gives the maximal subsystem dimension approximately equal to the number of subsystems. The reason for this is that it provides a good balance between having subsystems which are too large (and so susceptible to numerical problems) and ones which are too small (and so neglect appreciable cross-mode interaction). For the system studied here, such a choice can be seen to be $\rho_{ct} = 0.2$, giving 27 individual subsystems (some of which consist of single modes), and a maximum subsystem dimension of 17.

Figures 2-4 similarly illustrate the role of the overall correlation coefficient threshold ρ_{ot}. These figures can be seen to be broadly similar to Figure 1. The main difference is that the choice made for ρ_{ct} affects the initial, controllability-based, subsystem decomposition, which is then used to compute the final subsystem decomposition based on ρ_{ot}. Figures 2-4 demonstrate this by letting ρ_{ot} run through all possible values between 0 and 1 for ρ_{ct} equal to 0.1, 0.2 or 0.3, respectively. The differences between the three plots can clearly be seen: the main distinction is that the low controllability threshold value of 0.1 tends to give rise to larger final subsystems. This makes sense, as a lower threshold leads to more modes being considered sufficiently correlated that they must be grouped in the same subsystem. A rule-of-thumb for overall threshold that has generally been found to work well is simply to take $\rho_{ot} = \rho_{ct}$.

Table 1 then quantifies the effect of different choices for the reduced model order n_r. In all cases, the controllability and overall thresholds are taken to be equal, $\rho_{ot} = \rho_{ct} = \rho_t$, and this value is allowed to range over values between 0 (balancing) and 1 (modal

truncation). The numbers that are tabulated are the step response relative error norms, as defined by (4), for the resulting reduced-order models. It is significant to note that the new subsystem balancing technique gives better results than those of standard balancing or modal truncation for all the reduced orders considered. For instance, reducing the original 84-mode system to a 10-mode model by subsystem balancing with $\rho_t = 0.2$ gives rise to a relative error of less than 2%. By contrast, the error obtained for a model of the same dimension by means of balancing is over twice as large, and that produced by modal truncation is nearly three times as great.

TABLE 1.- STEP RESPONSE RELATIVE ERRORS

n_r	$\rho_t = 0$ Balanc	$\rho_t = .1$	$\rho_t = .2$	$\rho_t = .3$	$\rho_t = .4$	$\rho_t = .5$	$\rho_t = .6$	$\rho_t = .7$	$\rho_t = 1$ Modal
10	0.0416	0.0235	**0.0196**	**0.0196**	0.0218	0.0198	0.0233	0.0233	0.0556
11	0.0370	0.0226	0.0169	0.0169	0.0191	**0.0164**	0.0203	0.0203	0.0408
12	0.0188	0.0230	0.0161	0.0161	0.0138	**0.0117**	0.0203	0.0203	0.0179
13	0.0199	0.0207	0.0125	0.0125	0.0136	0.0117	**0.0082**	**0.0082**	0.0153
14	0.0096	0.0215	0.0129	0.0129	0.0145	**0.0057**	0.0058	0.0058	0.0080
15	0.0070	0.0180	0.0072	0.0108	0.0094	0.0071	**0.0050**	**0.0050**	0.0058
16	0.0073	0.0081	0.0067	0.0067	0.0091	0.0072	**0.0050**	0.0051	0.0058
17	0.0070	0.0081	0.0064	0.0063	0.0087	**0.0041**	**0.0041**	**0.0041**	0.0049
18	0.0049	0.0046	0.0038	0.0037	0.0042	**0.0035**	0.0036	0.0036	0.0046
19	0.0052	0.0050	0.0036	0.0036	0.0040	**0.0032**	0.0033	0.0033	0.0044
20	0.0044	0.0050	0.0042	0.0042	0.0041	0.0043	**0.0040**	**0.0040**	0.0052
21	0.0044	0.0052	0.0045	0.0045	0.0043	0.0045	**0.0036**	0.0037	0.0060
22	0.0043	0.0041	0.0035	0.0035	**0.0030**	0.0035	0.0040	0.0040	0.0079
23	0.0034	0.0034	0.0033	0.0033	**0.0029**	0.0032	0.0031	0.0031	0.0059
24	0.0038	0.0033	0.0026	0.0026	**0.0024**	0.0025	0.0032	0.0031	0.0039
25	0.0026	0.0020	0.0020	0.0020	0.0018	0.0018	**0.0016**	**0.0016**	0.0029

Figure 5 illustrates this further by plotting the step responses of both the full system and the 10-mode reduced-order model obtained for $\rho_t = 0.2$. It can be seen that the plots are nearly indistinguishable, despite the nearly order-of-magnitude reduction of system dimension. Figure 6 confirms this by plotting the step response error between the reduced model and the original system: note the different y-axis scale factors. For comparison, Figure 7 gives the error that would be obtained for a 10-mode model generated using modal truncation; again note the change in y-axis scale. The increase in this error over that obtained by subsystem balancing is clear. Finally, Figure 8 overlays the frequency response plots of the full system and the 10-mode reduced model obtained by subsystem balancing for $\rho_t = 0.2$. (These plots are weighted by a ω^{-1} term, so as to reflect the fact that a step input is being applied.) The close match between the two models is clearly visible in the frequency domain, just as it was in the corresponding time histories. It is also easy to identify which flexible modes are strongly excited by a prolonged, step-like, plume excitation force.

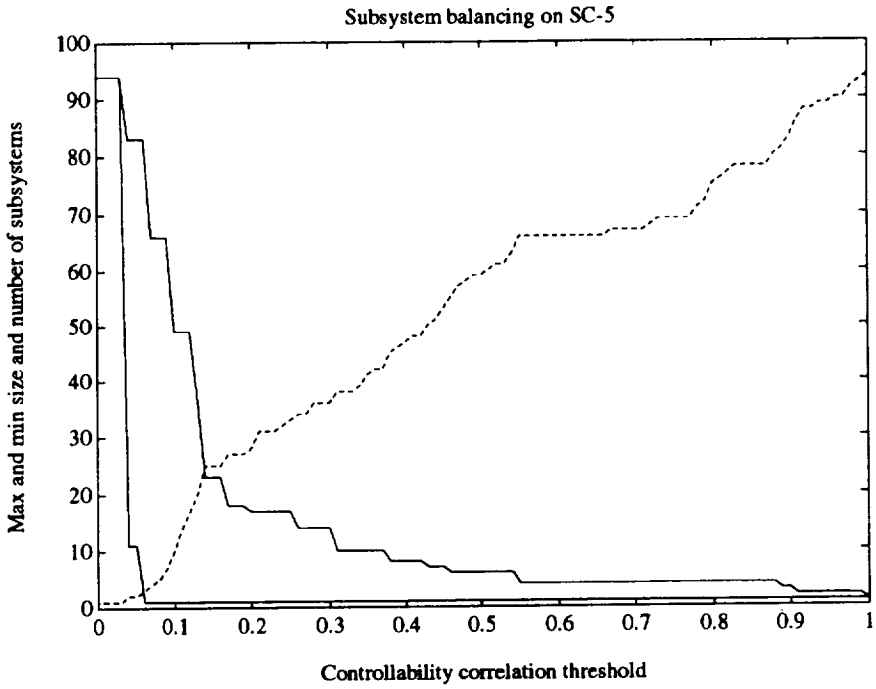


Figure 1.- Effect of varying controllability threshold level.

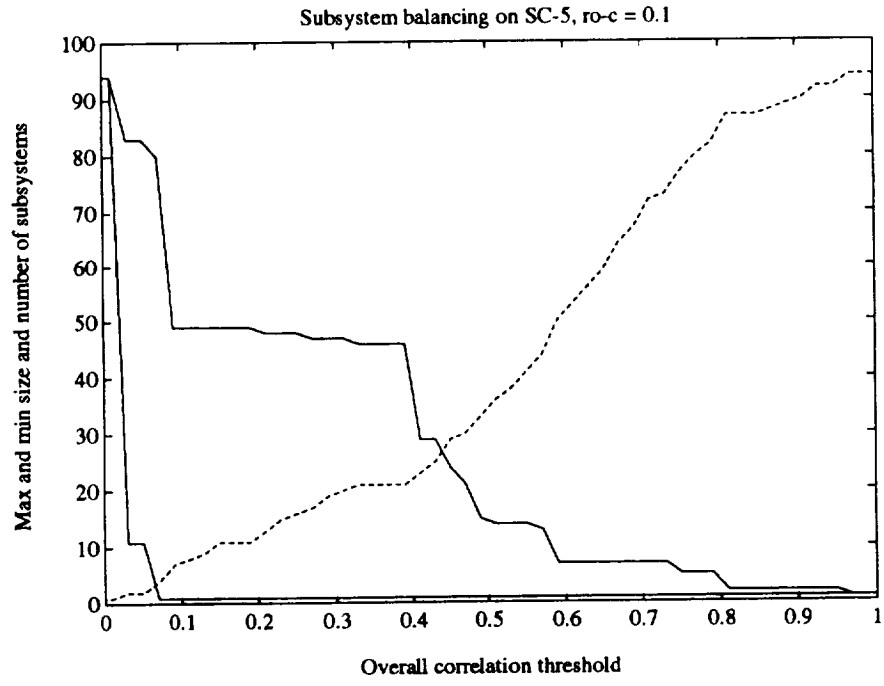


Figure 2.- Effect of varying overall threshold level for $\rho_{ct} = 0.1$.

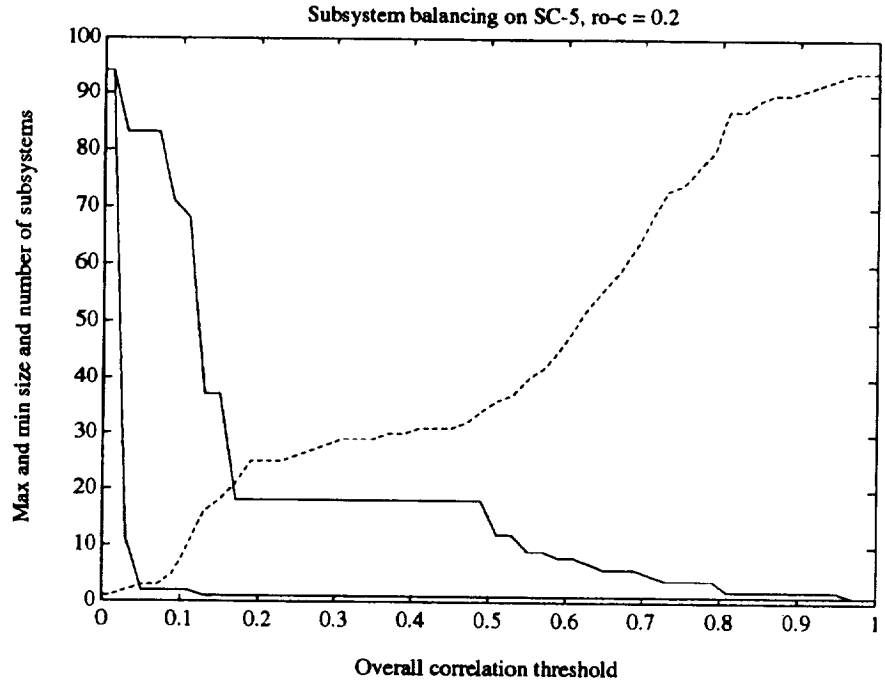


Figure 3.- Effect of varying overall threshold level for $\rho_{ct} = 0.2$.

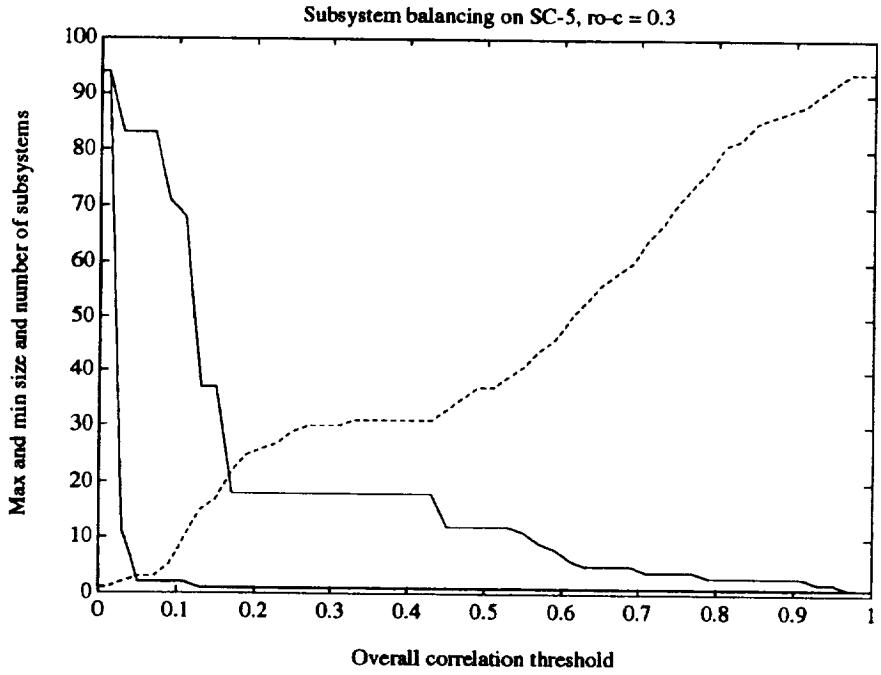


Figure 4.- Effect of varying overall threshold level for $\rho_{ct} = 0.3$.

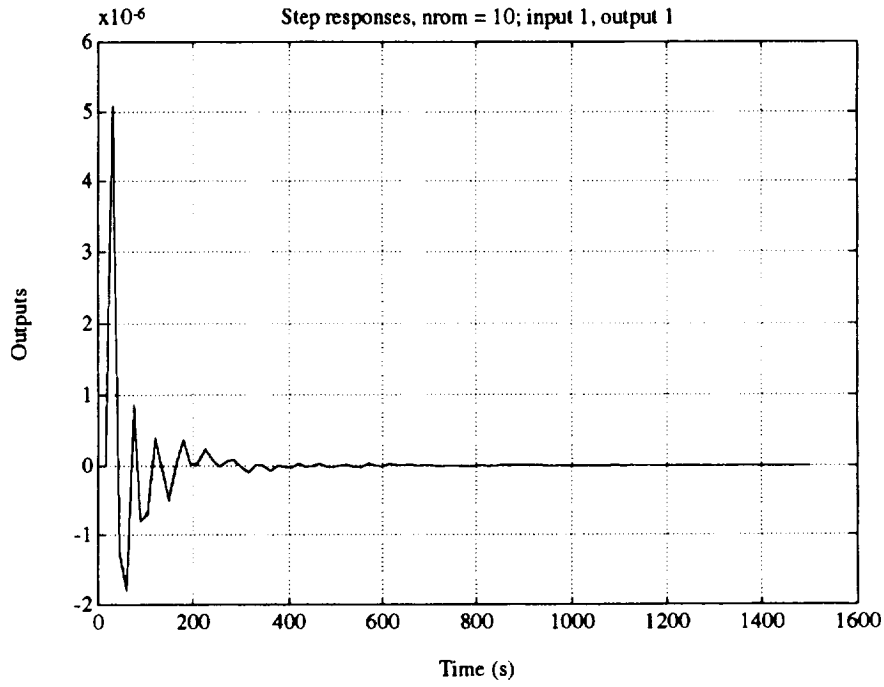


Figure 5.- Step responses of full and reduced models.

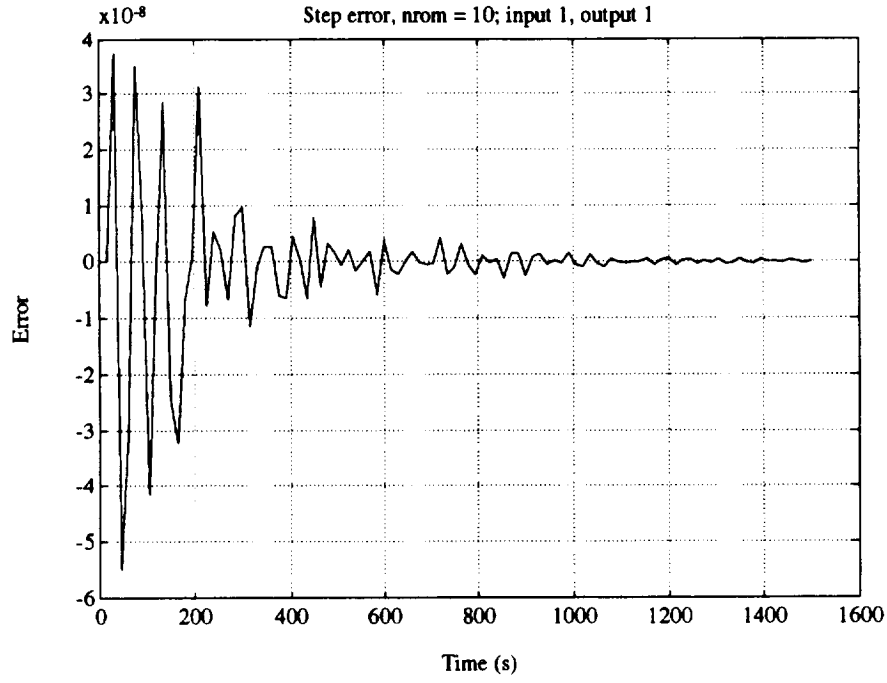


Figure 6.- Step response error of reduced model.

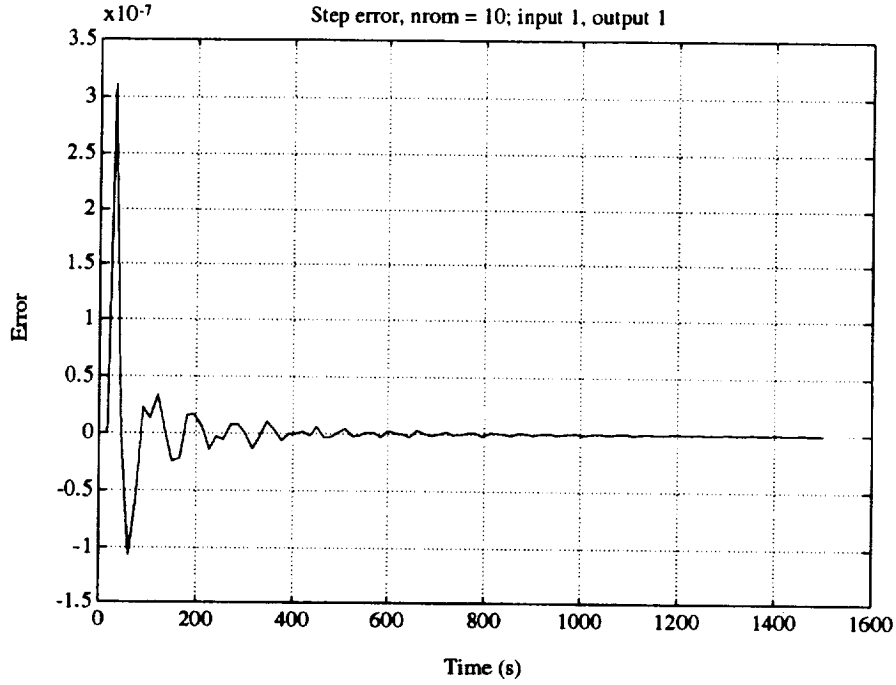


Figure 7.- Step response error of modal truncation model.

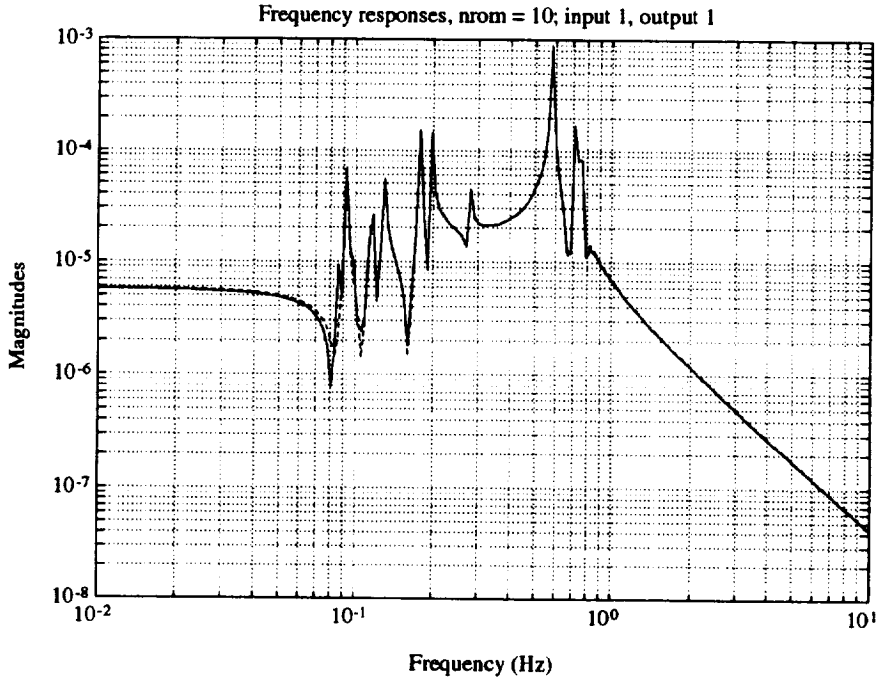


Figure 8.- Frequency responses of full and reduced models.

REFERENCES

1. R.R. Craig, *Structural Dynamics*, New York: Wiley, 1981.

2. B.C. Moore, 'Principal Component Analysis in Linear Systems: Controllability, Observability, and Model Reduction', *IEEE Transactions on Automatic Control*, Vol. 26, Feb. 1981, pp. 17-32.

3. T.W.C. Williams and W.K Gawronski, 'Model Reduction for Flexible Spacecraft with Clustered Natural Frequencies', invited paper, 3rd NASA/NSF/DoD Workshop on Aerospace Computational Control, Oxnard, CA, Aug. 1989.

4. W.K Gawronski and T.W.C. Williams, 'Model Reduction for Flexible Space Structures', *Journal of Guidance, Control, and Dynamics*, Vol. 14, Jan.-Feb. 1991, pp. 68-76.

5. T.W.C. Williams and M. Mostarshedi, 'Model Reduction Results for Flexible Space Structures', presented at 5th NASA/DoD Control-Structures Interaction Technology Conference, Lake Tahoe, Nevada, Mar. 1992.

6. T.W.C. Williams, 'Closed-Form Grammians and Model Reduction for Flexible Space Structures', *IEEE Transactions on Automatic Control*, Vol. 35, Mar. 1990, pp. 379-382.

7. M.J. Balas, 'Trends in Large Space Structure Control Theory: Fondest Hopes, Wildest Dreams', *IEEE Transactions on Automatic Control*, Vol. 27, 1982, pp. 522-535.

8. M.S. Tombs and I. Postlethwaite, 'Truncated Balanced Realization of a Stable Non-Minimal State-Space System', *International Journal of Control*, Vol. 46, 1987, pp. 1319-1330.

9. M.G. Safanov and R.Y. Chiang, 'A Schur Method for Balanced Model Reduction', *Proc. American Control Conference*, 1988, pp. 1036-1040.

10. R.E. Skelton and P. Kabamba, 'Comments on "Balanced Gains and Their Significance for L^2 Model Reduction"', *IEEE Transactions on Automatic Control*, Vol. 31, Aug. 1986, pp. 796-797.

11. H.T. Chang, 'SSF SC-4 Attitude Control Simulation', Dynacs Engineering Co. Report DYNACS-93-HM16, May 24, 1993.

12. T.W.C. Williams, 'Model Reduction for Space Station Freedom', NASA/ASEE Summer Faculty Fellowship Program Final Report, Aug. 1992.

N94-25384

A Multimedia Adult Literacy Package: Combining NASA Technology, Instructional Design Theory, and Authentic Literacy Concepts

Final Report
NASA/ASEE Summer Faculty Fellowship Program--1993
Johnson Space Center

Prepared By:	Jerry W. Willis, Ph.D.
Academic Rank:	Professor
University & Department:	University of Houston College of Education Department of Curriculum & Instruction Houston, Texas 77581
NASA/JSC	
Directorate:	Information Systems Directorate
Branch:	Software Technology Branch Advanced Software Architectures
JSC colleague:	James Villareal
Date Submitted:	August 27, 1993
Contract Number:	NGT-44-001-800

ABSTRACT

For a number of years, the Software Technology Branch of the Information Systems Directorate has been involved in the application of cutting edge hardware and software technologies to instructional tasks related to NASA projects. The Branch has developed intelligent computer aided training shells, instructional applications of virtual reality and multimedia, and computer-based instructional packages that use fuzzy logic for both instructional and diagnostic decision making. One outcome of the work on space-related technology-supported instruction has been the creation of a significant pool of human talent in the Branch with current expertise on the cutting edges of instructional technologies. When the human talent is combined with advanced technologies for graphics, sound, video, CD-ROM, and high speed computing, the result is a powerful research and development group that both contributes to the applied foundations of instructional technology and creates effective instructional packages that take advantage of a range of advanced technologies.

Several Branch projects are currently underway that combine NASA-developed expertise to significant instructional problems in public education. The Branch, for example, has developed intelligent computer aided software to help high school students learn physics and staff are currently working on a project to produce educational software for young children with language deficits. This report deals with another project, the Adult Literacy Tutor. Unfortunately, while there are a number of computer-based instructional packages available for adult literacy instruction most of them are based on the same instructional models that failed these students when they were in school. The teacher-centered, discrete skill and drill-oriented, instructional strategies, even when they are supported by color computer graphics and animation, that form the foundation for most of the computer-based literacy packages currently on the market may not be the most effective or most desirable way to use computer technology in literacy programs. This project is developing a series of instructional packages that are based on a different instructional model - authentic instruction. The instructional development model used to create these packages is also different. Instead of using the traditional 5 stage linear, sequential model based on behavioral learning theory, the project uses the Recursive, Reflective Design and Development Model (R2D2) that is based on cognitive learning theory, particularly the social constructivism of Vygotsky, and an epistemology based on critical theory.

Using alternative instructional and instructional development theories, the result of the summer faculty fellowship is LiteraCity, a multimedia adult literacy instructional package that is a simulation of finding and applying for a job. The program, which is about 120 megabytes, is distributed on CD-ROM.

Introduction

The Adult Literacy Project is a joint effort of NASA/JSC, the Department of Justice National Institute of Corrections, the University of Houston Center for Information Technology in Education, and the Houston Read Commission. Now in its third year, the project emerged from a long planning and basic research phase in October of 1993 and began a year of product development efforts. The work of the project is based on a major societal need.

There are, unfortunately, millions of adults in this country who are functionally illiterate. That is, they do not have the basic reading skills needed to successfully perform ordinary day-to-day living skills such as reading a bus schedule, completing a job application form, reading labels on food containers at the supermarket, or reading a daily newspaper. In some sections of the country the rate of adult illiteracy is as high as 35%. And, although adults become adept at hiding their inability to read, it prevents millions of people from full participation in society. Many fundamental aspects of citizenship in a democratic society -- from making informed voting decisions to holding a steady job -- increasingly require the ability to read.

This paper describes one component of a project to design, develop, and produce a series of computer-based, multimedia instructional packages. The packages are being developed for use in adult literacy programs, particularly in correctional education centers. They use the concepts of authentic instruction and authentic assessment to guide development. All the packages to be developed are instructional simulations. The first is a simulation of "finding a friend a job."

The Design Process

To develop the instructional package we used a heavily modified version of a traditional instructional development (ID) procedure. We adapted the Four D instructional development model of Thiagarajan, Semmel, and Semmel (1974), but it is very similar to models developed by Dick and Carey (1989) and many others. This model has four stages: (1) Define, (2) Design, (3) Develop, and (4) Disseminate. Although this model does provide a framework within which to organize and manage the process of developing instructional materials, it has a decidedly behavioral and linear flavor. Because the instructional packages we are developing are based on cognitive models of learning and teaching, the relatively poor "fit" between the ID model we had adopted and the process we were actually using quickly became obvious. Traditional ID models seem most suitable when the goal is to develop traditional, text-focused tutorials based on relatively standard behavioral learning theory. A project that strays from tutorials or from behavioral learning theory may find many aspects of the traditional ID models become barriers rather than facilitators to smooth, effective development of materials. In this project, the book, *Learning With Interactive Multimedia: Developing and Using Multimedia Tools in Education*, by Sueann Ambron and Kristina Hooper (Microsoft Press, 1990) was a major inspiration for modifying and transforming the ID model we were using.

In the following sections, each of the stages in the Four D model is summarized, with extensions and revisions related to the development of multimedia materials based on cognitive instructional theories. Most of the extensions and revisions reflect two major conceptual shifts in the ID model we used: recursion and reflection. While the 4 D Model is linear, we used it recursively which reflects one aspect of a more cognitive model of learning as well as of instructional development. The steps in the Four D model are described below in sequence that implies they are stages or phases in the ID process which must be completed one after the other. As we used the model, the 4 Ds were not steps to be completed one after the other as if they were layers of bricks in a wall. They were, instead, tasks to be completed. Unlike each layer of bricks in a wall, the tasks in the process may be addressed many times as the project progresses. Design and Development is thus a recursive rather than a linear process.

The other major change in the Four D model involves reflection in action (Schon, 1991). Schon's work on reflective practice emphasizes the need to think about and revise practice based on careful observation and analysis of what is happening in the practice environment. That approach, when applied to ID, places a heavy emphasis on obtaining feedback from students and instructors in the environment where the material will ultimately be used. In many ways, this represents an expansion of the formative evaluation procedures included in many behavioral ID models. The ID model we use in this project is thus a Recursive, Reflective Design and Development Model (R2D2).

The Definition Task

In the traditional ID model this task includes the preliminary steps required to develop any type of instructional material. The purpose of Definition subtasks is the identification and definition of instructional requirements. The five subtasks in this stage include (1) front-end analysis, (2) learner analysis, (3) task analysis, (4) concept analysis, and (5) specifying instructional objectives.

Front End Analysis. The initial step in developing instructional material is an evaluation of the need. This process generally involves a specification of the need and an evaluation of existing materials to determine if needs are already being met. The front-end analysis for this project indicated some computer-supported instructional materials were available for adult literacy programs, but almost all of them were drill and practice or tutorial packages designed from a distinctly behavioral perspective.

Learner Analysis. This step traditionally involves developing a clear understanding of the target students. In this case, the primary users are adult literacy students. We treated learner analysis as an on-going process rather than a preliminary phase of development. We worked with a number of students at the Houston Read Commission, but the format was not a traditional learner analysis that involves assessment of learner skills. Instead, the students became part of the development team. We asked for input on interesting themes for the simulation, discussed topics that should be in the simulation, and received feedback on visual

"look and feel" issues as well as instructional strategies the students preferred. Instructors at the Houston Read Commission also gave input on what their students needed, what they preferred, and ways the simulation could be made more interesting and relevant.

Task and Concept Analysis. After the need has been established and relevant student preferences, characteristics, and needs defined, the next step in a traditional ID model is to define the major skills to be acquired by the students and to analyze those skills. In traditional behaviorally oriented projects the focus is on breaking things down into subunits or components that can be taught separately. In reading, this approach leads to a discrete skills model of instruction that targets things like short vowel sounds and consonant digraphs. The result of this model in adult literacy programs is often relatively boring exercises that attempt to remediate weak underlying skills on the assumption that when the student develops those cognitive "muscles" reading will improve because reading is really a collection of skills. This model has not served adult literacy well and it may be one reason for the high dropout rate in many programs as well as the lack of enthusiasm shown by many participants who truly want to learn to read. In addition, many adults who do not have a functional level of reading ability have already completed and/or failed at many such exercises during their school years. Many adults who do not have functional literacy skills, but many do have significant negative associations with them.

Using the R2D2 Model we approached task and concept analysis from a somewhat different perspective: authentic assessment and instruction. The primary goal was to deliver instruction and provide assessment directly related to authentic tasks an adult in this culture completes by reading. Those tasks are the focal point of the project and both instruction and assessment focuses on them. Thus, during instruction the task of the reader will be to accomplish a common task such as finding and applying for a job, locating and renting an apartment, using social services, or buying groceries. Assessment focuses on aspects of the reader's approach to the task that limit or prevent success. That is not to say that no discrete skills analysis or skills instruction is conducted or used. It simply means that it is framed within the context of how it relates to a particular task that is real or *authentic* for the student.

When task and concept analysis is considered within the framework of authentic assessment and authentic instruction, the analysis becomes less a job to be done once and for all at the beginning of the development process and more a continuous aspect of ID. For the project described in this paper the task was "Finding a Friend a Job" which involved selecting appropriate jobs, obtaining and completing the application, and submitting the application to the potential employer. The instructional product was built around the task. Actual instruction and help learning to read is available to the student when, and only when, the student requests it. The instruction and help given is determined by the student who selects it in order to complete the authentic task at hand.

Specifying Instructional Objectives. This is normally the final step in the first stage of a traditional ID model. It involves converting the results of the analysis of

tasks and concepts into a set of objectives. They in turn provide the basis for designing the instructional package and developing evaluation and assessment strategies. As with task and concept analysis, specifying instructional objectives changes somewhat when viewed from the perspective of the R2D2 Model. In behavioral theory, specific instructional objectives such as "When the reader is faced with a paragraph of directions on how to drive from home to a potential employer's office, he or she draws a correct map based on those directions" are needed to guide the design process. In the R2D2 Model, which is based on authentic instruction and assessment, specific objectives evolve naturally out of the creation of a realistic instructional environment that represents a real world literacy task. Objectives keep designers focused and help ensure that the instruction leads to learning that transfers to the real world. It is not as important to write specific objectives first when the design involves creating an instructional version of some aspect of the real world. Concepts such as information landscapes, user support, and user interface become much more important than the development of a long list of specific instructional objectives.

The Design Task

The focus of this task is the design of prototype instructional material. Two subtasks in most traditional ID models, Media Selection and Format Selection, are completed along with a third, Selection of a Development Environment. Two other traditional subtasks, Initial Design and Evaluation Strategy, are folded into other tasks rather than treated separately.

Media and Format Selection. The media we selected was interactive multimedia and the format selected was simulation. Because of the type of computer and media resources available, or affordable, in most adult literacy programs, we selected a single-screen simulation that can be run on a Macintosh or IBM computer with a CD-ROM drive and a sound system. The package puts the reader in the primary role in a simulation of an interesting and typical real world activity - finding a job. During the prototype development stage, the entire package, including sound effects, music, and speech files were stored in files on the computer's hard disk drive. The final version will be distributed on a CD-ROM that includes the computer program in track 1. All the graphics are included in track 1 as part of the program. The audio, which includes spoken directions, help, and dialog from characters in the simulation, is currently in a standard computer sound format. One alternative is to put the sound in CD-Audio format on tracks 2 through 99 of the CD-ROM.

The "Finding a Friend a Job" scenario used in the first instructional package requires the reader to select a friend, decide where the friend should look for a job, select a job from those advertised that fits the friend's needs and abilities, and correctly complete a three part job application that serves as a model for the friend. If the job selected matches the needs and talents of the friend, the simulation ends with a call from the friend who says "I got the job .... and I could not have done it without your help." If the job selected is inappropriate the friend calls to tell the reader there

was a problem and asks for more help. The simulation can then begin again.

Throughout the simulation, the reader can switch from simulation mode to help mode by clicking a button. When the reader does not know a word or phrase, he or she can ask for assistance. Readers click the problem word or words and then choose which of five different reading strategies they think would be most helpful. The five strategies, which are represented by 5 icons at the bottom of the help screens, are phonics, syntax, context, pronounce for the student, and read and explain. For each strategy selected readers apply it with the guidance of spoken directions. Throughout the simulation, readers can ask for help by clicking a question mark button and replay verbal directions and explanations by clicking a speaker button on the top right of the screen. They can replay the voice version of printed directions, explanations, and information by clicking small speaker icons displayed beside the text. Spoken help, directions, and assistance is available in either Spanish or English.

Selecting a Development Environment. The core of the development environment was Authorware Professional on the Macintosh (APM) supplemented by C language subroutines as well as sound and graphics software including Photoshop, SoundEdit, SoundWave, and Canvas. Graphics, music clips, and sound effects were used from a number of clip collections including Killer Tracks music clips and Wraptures backgrounds. A major advantage of APM as the development platform is its ability to support a wide range of utility and support software and hardware. APM has provisions for integrating sound including music clips and voice from many sources, still images and video from CD-ROM and laser discs, and graphics from many paint and draw programs.

Authorware Professional for the Macintosh (APM) is widely recognized as a powerful development system. It is used by a number of NASA subcontractors, including Rockwell, and by Boeing. APM has a number of advantages over developing instructional packages in standard programming languages such as C. Perhaps the most important advantage for this project is the ease of learning factor. Experienced instructional designers can be "up and running" in Authorware in less than a week while estimates of the time it takes someone to become a proficient independent C programmer are as high as two to three years of full-time work and study. In addition, products developed on APM can be transferred to the Authorware Professional version for Windows (APW). APM's ease of learning, and ease of use, facilitates an important aspect of the development team's work pattern. In traditional ID models, each person takes a specialist's role. For example, a graphic artist may create illustrations, an instructional designer creates sequences or frames of instruction, and a content expert either creates or approves of the content included in the package. In the R2D2 Model as used in this project, the team was composed of "specialists" who were primarily responsible for different aspects of the program such as graphics or sound. One member of the team, however, served as the designated programmer. The roles were, not, however, watertight compartments. Instead, almost all the team members worked some on all aspects of development. The content specialist, for example, wrote and revised some of the Authorware code and redesigned some of the

icons. The sound specialist wrote some of the script and designed several of the help and instructional strategies. APM's ease of use made it possible for every team member to program. The interactive nature of APM made design recursion possible. Without a development environment like APM, reflection and recursion would be less useful, or discouragingly useless, because it is too difficult to make changes and adjustments to the program once it is "set in code." For example, instead of storyboarding an instructional sequence and then turning it over to the programmer to be written, the team could quickly answer many "what if" questions. What if we change the location of the icons, what if we add an explanatory screen between two existing screens, what if we change the layers a reader moves through to get help with an unknown word? In APM it is easy to set "Start" and "Stop" flags at points in the program, make changes in that section, and then run only that section to see how it will appear to students. The same feature makes it very easy to get student and teacher input and suggestions on different components as the program is evolving. For example, the team created four different "look and feel" designs and asked sixty adult literacy students to select the one they preferred. The same group listened to eight different musical styles and selected the one the liked the best. Such on-going and recursive feedback-revise cycles are much easier to complete using APM than regular programming languages.

APM is not a perfect development environment, but it has much to offer when the ID model is recursive and reflective. An educational product is as much an artistic creation as it is a technical creation. We are all aware of technically correct educational materials, manuals, and documentation that numb the mind and discourage their use. This project aims to create products that appeal to the user, that have the snap and crackle that comes from well placed and well designed graphics, products that have an appealing premise creatively executed. A traditional, top down, linear model of development is unlikely to produce such a product. What is needed is a development model that is the equivalent to the "management by walking around" movement in corporate leadership. That is, the designers and developers need to be able to "walk around" the program and try out possible alternative designs and arrangements with regular and ongoing feedback from potential users and consumers. The approach is easily implemented with Authorware. Changes, even major changes, are relatively easy to make while your program is running. And changes, once made, can be run immediately to view the effect from the perspective of the student who will be using the package. This advantage, which avoids the designer-to-programmer-to-software-to-designer cycle that has doomed so many instructional development projects, allows developers to "play with" many different alternatives in much the same way a musician can play with changes in a musical composition. This non-linear, real time approach supports improvisation, which leads to exciting designs. Eric Holsinger, in his article titled "Nonlinear systems enhance video editing" (MacWeek, 10, 26, 1992) makes the same point about electronic video editing. "Unlike standard video-editing systems, an editor can quickly cut and paste sections of the presentation together in any order. A user can save and compare different edits of the presentation

easily, whereas starting a new version on a traditional off-line system requires you to start the edit from scratch each time."

Initial Design and Evaluation Strategy. The final two steps in the traditional Design stage are Initial Design and construction of an Evaluation Strategy. Because this project involves a product based on authentic instruction and authentic assessment theory there is less concern about a separate 'test" of success. The simulation has a built-in assessment component and there is much less need for an external evaluation strategy such as pre- and posttests. The focus was thus on Initial Design. We approached the task of creating an instructional simulation by first developing an overall concept. Then we worked concurrently on three aspects of the general format and design:

1. Surface design - screen layout, typography, language, graphics, illustrations, sound.
2. Interface design - look and feel, user interaction, help, support, navigation, metaphors.
3. Scenario - sequence of simulation, options/choices, results.

We approached these three components by defining one path through the simulation that included every type of interaction, every type of help, and every type of outcome and feedback. We then completed all the work on that path so that the simulation ran from start to finish as planned - if one particular path was selected. When we completed the surface design, interface design, and scenario for one path, students and teachers critiqued the program. We modified the program based on that feedback and used the Authorware components in this path to build all the other paths. This reduced the effort required to develop the full simulation. One path serves as a template for the other paths through the simulation. In addition, many components in the Finding a Friend a Job simulation will be reused in the next simulation in the LiteraCity series.

The Development Task

As the Design Tasks are completed, the Development Tasks become more and more important and relevant. In the R2D2 model, that does not mean, however, that development tasks are completed only after all the design tasks have been done. In practice, design and development tasks are completed concurrently. An initial prototype is collaboratively developed and evaluated. Collaborative development involves the entire design team as well as consultants such as front line practitioners with recognized expertise and experience. As components of the prototype were developed they were subjected to two types of evaluation: expert appraisal and developmental tryouts. Both provide feedback that will be used to revise the prototype. In this project expert appraisal of the package was obtained from adult literacy practitioners and correctional education specialists. Developmental testing was carried out in Houston-area adult literacy programs.

We currently have a working prototype developed and, after a few more develop-get feedback-revise cycles we will concentrate on interesting an educational software

publisher in distributing the product. Because Authorware is a flexible development environment it is easy to make changes based on the feedback from both expert appraisal and developmental tryouts.

Dissemination

Once a package has been revised based on feedback from expert appraisals and developmental tryouts, the final product can be subjected to a summative evaluation. The summative evaluation generally involves using the package with a fresh group of students in an environment like the one where the material is to be used. Qualitative and quantitative data gathered from this evaluation is often included in the final teacher's manual included with the package. Few educational software products are subjected to a summative evaluation and when the development work to this point has included extensive feedback from teachers as well as tryouts with students, a summative evaluation may not be as important as it would be if the product had been developed without extensive input from the field. A summative evaluation will not be conducted on the literacy simulations.

The final three Dissemination subtasks are Final Packaging, Diffusion, and Adoption. Final packaging involves creating and producing any necessary print materials (such as student guides and teacher's manuals), creating a master for the CD-ROM or laser disc and pressing a quantity of CD-ROMs for initial use. If we are able to attract a commercial partner before final packaging, some of the work in this phase will be handed off to the partner. The commercial partner will also carry the primary responsibility for diffusion and adoption.

Some Comments on the Design and Development Process

For this project a design team was organized. A wide range of specialists were required including programmers, instructional design specialists, content specialists and instructors, computer graphic artists, and project management leaders. The development of an effective team is not an easy task. There are, in fact, probably many more failures than successes in this field. The difficulty of the task is illustrated by Mark Heyer's article in the Microsoft Press book, *CD-ROM: The New Papyrus*. In his chapter Heyer, who has worked with Sony and Group W Westinghouse on CD-ROM and laser disc projects, describes what is hopefully a worst case scenario:

> The conception, design, and production of visual/computer interactive programming is the newest challenge facing the creative community. . . .
>
> In the videodisc industry, some organizations understand this problem [getting designers, visual producers, and programmers to work together] and have begun to create teams. In most cases, though, members of the three groups are separated in space and time. A few companies have even mandated that programmers will not talk to designers or video producers -- a clear failure mode.

> Generally the process follows a path something like this: First a design document is generated and argued about for four to six months. So far it's a paper war. At best, 10 percent of the people who are judging and modifying the paper concept will actually understand the nature of interactive video.
>
> Next the video is produced. Whether the content is motion or stills, this step normally consumes 80 percent of the total project time and budget. . . .
>
> Then, after the visual material is mastered on a video or data tape, the control software is needed. In the case of videodiscs, most of which use external computer control, the disc is pressed and then the control software is written to conform with what has been put on the discs. Sadly, in many cases the programmer isn't even hired until after the disc is done. At that point the opportunity to make changes has passed. . . .
>
> CD-ROM has an extra burden to consider, since the control software has to be pressed on the disc along with the visuals. Everything has to be perfect the moment the master is made. In reality, many producers will bear the expense of pressing discs with just the visuals as an aid to programming. This works about as well as any batch processing scheme, but it still doesn't allow for many visual change cycles, or for creative design input during programming. The computer programmer works with a fully interactive computer programming language and attempts to control a videodisc, which can be done in many ways, but the read-only pressed videodisc is in fact fixed and unchangeable.
>
> The challenge of understanding these problems, and to some extent working around them, has occupied a whole generation of videodisc producers. We are just now in possession of design tools and methods which dramatically cut the development time and costs for videodiscs. The same techniques and equipment will work for CD-ROM. (Heyer, 1987, pp. 352-353)

Heyer paints a fairly accurate picture of many development efforts. Fortunately, he also offers some very useful suggestions. He suggests, for example, that design tools and systems must be developed and used that are relatively inexpensive and easy to use. Perhaps the most important point he made relative to this project relates to what he calls Interactive Editing. "For creative artists and designers, whether visual or computer, the ability to make unlimited small changes is an absolute

benefit." Heyer suggests using a videodisc emulator so that "designers, visual artists, and programmers are working together on program creation in real time. Hundreds of changes can be made in a day, and each change is instantly demonstrable. Design and development time are cut to a fraction of what they used to be." Just such a recursive, real-time development process is essential in a project that depends on creativity for success. Heyer proposes that the design process begin with a crude storyboard that may be done on paper or on the computer using rough video and low fidelity graphics. Team members can work on the rough storyboard that may be done on paper or on the computer using rough video and low fidelity graphics. Team members can work on the rough storyboard, which is really a rough prototype, and experiment with a wide range of design options and alternatives. As the prototype matures, "real video and graphics are substituted for storyboard images, the code is finalized, and the project is finally sent off for mastering." Heyer sees three major advantages in this approach. Risk is reduced because fundamental problems can be discovered early. Because design and programming occurs at the same time, there is less likelihood that the team will commit to a design that cannot be executed. A second advantage is the encouragement of experimentation and revision which will enhance the quality of the final product. Finally the process of development becomes understandable and "each person knows exactly what the others are doing."

In this project a fully functioning path was developed through the entire simulation. Then the appropriate graphics, sound, and images were created and added to the program. That path was then evaluated by literacy teachers, literacy students, and experts in relevant areas such as graphics, instructional design, and assessment. Finally, the pattern and format that emerged from work on the path was extended to the other paths through the program.

A design and development approach that is recursive and allows for changes and revisions in "real time" calls for a collaborative team with a wide range of skills and expertise. It also calls for a team that can tolerate almost continuous change, adjustment, and fine tuning as well as the frustration that is probably inevitable when some work is abandoned or replaced by new material. The major drawback to traditional linear models of instructional development may be that they eliminate many opportunities for fine tuning and artistic enhancement. The major drawback to recursive models may be that everything is so fluid that team members have difficultly seeing progress and feeling good about their accomplishments to date. If everything can be changed at any stage of development, nothing is ever truly "finished" until the end of the project. Many team members need the feeling of closure that comes from finishing different stages of a long, complex project and teams that use a recursive model need to attend to those needs. Recursive approaches, like linear models, can be taken to extremes that are both frustrating and nonproductive. Whistler, for example, had great difficulty deciding when a painting was truly "finished." On several occasions he actually went to the homes of patrons who had purchased a painting and made small changes to it!

Non-linear, recursive development models must also deal with the fact that time, a factor in all projects, is linear. Almost all development projects have tight timelines, and recursive approaches call for nimble but energetic project management, particularly when the work of many different team members feeds into, and is required for, the work of others. Common problems in projects that use a linear development model -- assignments that are not completed on time, work that does not "fit" with work done by other team members, concepts and design formats that are changed by one group or individual but not communicated to others, and the "just a few more changes and it will be ready" phenomenon - can happen more often in recursive models and multiply the possibilities for problems, delays, and conflicts. Effective time management, long and short range project planning, successful team building, and strong overall project management are critical when a non-linear, recursive approach is used.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	Dec/93	Contractor Report

4. TITLE AND SUBTITLE
National Aeronautics and Space Administration (NASA)/American Society for Engineering Education (ASEE) Summer Faculty Fellowship Program - 1993, Volumes 1 and 2

5. FUNDING NUMBERS

6. AUTHOR(S)
William A. Hyman* and Stanley H. Goldstein, Editors

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Johnson Space Center
Houston, Texas 77058

8. PERFORMING ORGANIZATION REPORT NUMBERS

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, D.C. 20546-0001

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
CR 188271

11. SUPPLEMENTARY NOTES
* Texas A&M University, College Station, Texas

12a. DISTRIBUTION/AVAILABILITY STATEMENT
unclassified/unlimited
Available from the National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161 (703) 487-4600.

Subject Category: 99

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The JSC NASA/ASEE Summer Faculty Fellowship Program was conducted by Texas A&M University and JSC. The objectives of the program, which began nationally in 1964 and at JSC in 1965, are (1) to further the professional knowledge of qualified engineering and science faculty members; (2) to stimulate an exchange of ideas between participants and NASA; (3) to enrich and refresh the research and teaching activities of participants' institutions; and (4) to contribute to the research objectives of the NASA centers. Each faculty fellow spent at least 10 weeks at JSC engaged in a research project in collaboration with a NASA/JSC colleague. This document is a compilation of the final reports on the research projects completed by the faculty fellows during the summer of 1993.

14. SUBJECT TERMS
information transfer, research , research projects, urban research, engineering, science, universities, university program

15. NUMBER OF PAGES
488

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	UL